ENRICO DI BENEDETTO GIOVANNI CUSIMANO

APPUNTI DI

ISTITUZIONI DI METODI MATEMATICI PER LA FISICA

Appunti di *Istituzioni di metodi matematici per la Fisica*

Questo lavoro è basato sul corso tenuto durante l'A.A. 2023/24 dal Prof. Fabio Bagarello nell'ambito del Corso di Laurea in Scienze Fisiche erogato dall'Università degli Studi di Palermo.

La presente edizione si nutre degli appunti precedentemente redatti dal dott. Enrico Di Benedetto durante l'A.A. 2020/21. Una revisione si è ritenuta necessaria in luce degli immancabili errori presenti in ogni testo, delle modifiche apportate al corso durante gli anni e dalla voglia di migliorare la veste grafica del testo. Tale revisione è curata da Giovanni Cusimano e da Enrico Di Benedetto.

Questo lavoro non vuole sostituirsi ai testi consigliati nella scheda di trasparenza del corso; vorrebbe rappresentare un supporto allo studio senza alcuna pretesa di completezza. Gli autori, pertanto, consigliano fortemente l'utilizzo dei libri di testo in abbinato ai seguenti appunti, che rappresentano un punto di partenza.

Si ringraziano i numerosi colleghi che negli anni hanno usato queste note e soprattutto coloro che hanno segnalato con pazienza immancabili sviste e mancanze. I futuri lettori sono caldamente invitati a fare altrettanto.

We cordially welcome recommendations for enhancements or notifications of any errata contained within the subsequent text, which can be transmitted to the designated email addresses, specifically *enrico.dibenedetto@unipa.it* and *giovanni.cusimano.pi@gmail.com*.

We, the writers, fervently anticipate that these annotations may bestow invaluable support upon those who encounter challenges with the topics expounded upon in the course. *If you love this product, please consider donating.*

Respectfully,

Enrico Di Benedetto
Giovanni Cusimano

Indice

Capitolo 1

Spazi di Hilbert

In questo capitolo viene introdotta la nozione di *spazio di Hilbert*, la quale fornisce la base matematica di diverse teorie di grande importanza in fisica, tra cui la meccanica quantistica. Come complemento a tale introduzione, viene brevemente affrontato il tema dei *polinomi ortogonali* e vengono elencate alcune famiglie notevoli di polinomi.

1.1 Cos'è uno spazio di Hilbert?

Ricordiamo che uno *spazio vettoriale lineare* su campo $\mathbb{K}$ è un insieme di elementi, detti *vettori*, che soddisfa le proprietà di linearità e di chiusura rispetto alla somma e al prodotto per uno scalare scelto nel campo $\mathbb{K}$. Questo significa che ogni combinazione lineare di vettori dello spazio produce un vettore facente parte dello spazio, così come il prodotto di un vettore per uno scalare facente parte del campo su cui lo spazio è definito (tipicamente $\mathbb{R}$ o $\mathbb{C}$).

Avendo in mente questa definizione, possiamo introdurre gli spazi di Hilbert. Per comodità di notazione, specializziamo la nostra definizione al caso in cui il campo è l'insieme dei numeri complessi $\mathbb{C}$.

Definizione 1.1. *Uno* **spazio di Hilbert** *H è uno spazio dotato delle seguenti caratteristiche:*

1. *è uno spazio vettoriale lineare su campo $\mathbb{C}$,*
2. *è definita un'operazione binaria $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$, detta* **prodotto scalare**, *tale che*
   - *a)* $\langle f, g \rangle = \overline{\langle g, f \rangle}$,
   - *b)* $\langle f, \alpha g + \beta h \rangle = \alpha \langle f, g \rangle + \beta \langle f, h \rangle$ (**linearità nella seconda componente**),
   - *c)* $\mathbb{R} \ni \langle f, f \rangle \geq 0 \; \forall f \in H$,
   - *d)* $\langle f, f \rangle = 0 \Leftrightarrow f = 0$.
3. *è completo rispetto alla metrica indotta dal prodotto scalare, ovvero ogni successione di Cauchy converge ad un elemento di H.*

Nella proprietà (2c) è superfluo specificare che $\langle f, f \rangle \in \mathbb{R}$, dato che in $\mathbb{C}$ non esiste un ordinamento totale, ovvero $\mathbb{C}$ non è un campo totalmente ordinato. La proprietà (2d) si può rilassare notando che, ponendo $f = 0$, si ha

$$\langle 0, 0 \rangle = \langle 0, h - h \rangle = \langle 0, h \rangle - \langle 0, h \rangle = 0 \;\; \forall h \in H \tag{1.1}$$

da cui segue che l'implicazione $\Leftarrow$ in (2d) è sempre banalmente verificata imponendo l'implicazione $\Rightarrow$. La proprietà corrispondente a (2b) per la prima componente è

$$\langle \alpha f + \beta g, h \rangle = \overline{\langle h, \alpha f + \beta g \rangle} = \overline{\alpha \langle h, f \rangle + \beta \langle h, g \rangle} = \overline{\alpha} \langle f, h \rangle + \overline{\beta} \langle g, h \rangle . \tag{1.2}$$

Tale espressione fornisce una legge di **antilinearità nella prima componente**.

Esempio 1.1. Nel seguito, useremo la nozione di spazio di funzioni a potenza p-esima integrabile, comunemente indicato con $\mathscr{L}^p(D)$. Tale spazio contiene tutte le funzioni aventi come dominio D e Lebesgue-misurabili[a] tali per cui

$$\int_D |f(x)|^p \mathrm{d}x < \infty \quad \forall\, 1 \leq p \leq +\infty.$$

È un esercizio istruttivo verificare che tale spazio è uno spazio vettoriale lineare. Più complicato, ma ugualmente interessante, è dimostrare che tali spazi sono *spazi di Banach*, ovvero spazi metrici completi, indipendentemente dal valore di p.
Allo stesso modo, si introducono gli spazi delle successioni a potenza p-esima sommabile, indicati con $\ell^p(D)$, come gli spazi delle successioni degli $a_i \in D$ tali per cui

$$\sum_{i=1}^{\infty} |a_i|^p < \infty. \tag{1.3}$$

[a] v. Appendice B.

Esempio 1.2. È immediato verificare che i seguenti sono spazi di Hilbert:

1. l'insieme dei vettori ad n componenti reali $\mathbb{R}^n$, dotato del prodotto scalare $\langle a, b \rangle = \sum_{i=1}^n a_i b_i$,
2. l'insieme dei vettori ad n componenti complesse $\mathbb{C}^n$, dotato del prodotto scalare $\langle a, b \rangle = \sum_{i=1}^n \overline{a_i} b_i$,
3. l'insieme delle funzioni a quadrato sommabile $\mathscr{L}^2(\mathbb{R})$ dotato del prodotto scalare $\langle f, g \rangle = \int_{\mathbb{R}} \overline{f(x)} g(x)\, \mathrm{d}x$,
4. l'insieme delle successioni a quadrato sommabile $\ell^2(\mathbb{R})$ dotato del prodotto scalare $\langle a, b \rangle = \sum_{i=1}^{\infty} \overline{a_i} b_i$.

1.1.1 Norma e metrica indotta dal prodotto scalare in H

Una delle proprietà coinvolte nella definizione di uno spazio di Hilbert è la completezza rispetto alla metrica indotta dal prodotto scalare. Cerchiamo di fare chiarezza su come ricavare tale metrica.

La proprietà (2c) ci permette di introdurre il modulo o *norma* di un elemento dello spazio di Hilbert H a partire dal prodotto scalare stesso.

Definizione 1.2. *Chiamiamo* norma indotta *dal prodotto scalare la funzione*

$$\|f\| = \sqrt{\langle f, f \rangle} \tag{1.4}$$

È immediato verificare che, in conseguenza delle proprietà del prodotto scalare, tale espressione effettivamente rappresenti una norma in H, le cui proprietà per definizione sono:

1. $\|f\| \geq 0$ e $\|f\| = 0 \iff f = 0$,
2. $\|\alpha f\| = |\alpha|\|f\|$,
3. $\|f+g\| \leq \|f\| + \|g\|$.

Esempio 1.3. Consideriamo lo spazio delle funzioni continue su un compatto $[a,b]$ e definiamo la *norma uniforme* [a]

$$\|f(x)\|_U := \max_{x\in[a,b]} |f(x)|, \quad \forall f(x) \in C^0([a,b]). \tag{1.5}$$

Mostriamo per esercizio che tale espressione rappresenta effettivamente una norma in $C^0([a,b])$. Infatti

1. $\|\cdot\|_U \geq 0$ poiché $\|\cdot\|_U$ è definita come il valore massimo di una funzione non negativa. Inoltre, se $\|f\|_U = 0 \Rightarrow 0 \leq |f(x)| \leq 0 \forall x \in [a,b]$, che implica chiaramente $f(x) = 0$;
2. $\|\alpha f\|_U = \max\limits_{x\in[a,b]} |\alpha f| = |\alpha| \max\limits_{x\in[a,b]} |f| = |\alpha|\|f\|_U$;
3. $\|f+g\|_U = \max\limits_{x\in[a,b]} |f+g| \leq \max\limits_{x\in[a,b]} (|f|+|g|) = \|f\|_U + \|g\|_U$.

[a] Il nome di questa norma deriva dal fatto che la convergenza uniforme di una successione di funzioni è del tutto equivalente alla convergenza in tale norma.

L'introduzione di una norma indotta dal prodotto scalare rende ogni spazio di Hilbert uno spazio metrico, che risulta essere per definizione di spazio di Hilbert anche completo.

Precisiamo che, in generale, non tutti i prodotti scalari inducono una norma. Ciò equivale a dire che non tutti gli spazi metrici completi sono anche spazi di Hilbert. Il ragionamento in questo caso è molto simile a quanto si osserva nel rapporto tra spazi metrici e spazi normati, in cui non è difficile trovare metriche non indotte da alcuna norma, ovvero spazi metrici non dotati di alcuna norma.

Esempio 1.4. Due esempi di spazi normati che **non** sono spazi di Hilbert sono

1. $\mathscr{L}^1(D)$, dotato della norma $\|f\|_1 = \int_D |f(x)|\, dx$, ove $D \subseteq \mathbb{R}$;
2. $\ell^2(\mathbb{N})$, dotato della norma $\|a\|_\infty = \sup\limits_{j\in\mathbb{N}} |a_j|$.

Notiamo che in quest'ultimo caso è la scelta della norma infinito a impedire che questo spazio sia uno spazio di Hilbert. Esiste infatti una norma indotta dal prodotto scalare, ovvero $\|\cdot\|_2$, che rende $\ell^2(\mathbb{N})$ uno spazio di Hilbert.

In generale, il fatto che una norma sia indotta (o meno) da un prodotto scalare può essere abbastanza complicato da dimostrare, soprattutto se si adotta la strategia di *procedere per tentativi* nel costruire un prodotto scalare ad hoc. Una condizione necessaria affinché una norma sia indotta da un prodotto scalare è data dall'identità del parallelogramma.

Lemma 1.1 (Identità del parallelogramma). *Se la norma $\|\cdot\|$ è indotta da un prodotto scalare $\langle\cdot,\cdot\rangle$ avente le proprietà di linearità (2), si ha che*

$$\|f+g\|^2 + \|f-g\|^2 = 2\left(\|f\|^2 + \|g\|^2\right) \quad \forall f,g \in H. \tag{1.6}$$

Dimostrazione. Ricordando la Definizione 1.2, possiamo riscrivere il primo membro dell'espressione come

$$\begin{aligned}\|f+g\|^2+\|f-g\|^2 &= \langle f+g,f+g\rangle + \langle f-g,f-g\rangle \\ &= \langle f,f\rangle \cancel{+\langle f,g\rangle} \cancel{+\langle g,f\rangle} + \langle g,g\rangle + \langle f,f\rangle \cancel{-\langle f,g\rangle} \cancel{-\langle g,f\rangle} + \langle g,g\rangle \\ &= 2(\langle f,f\rangle + \langle g,g\rangle) = 2\left(\|f\|^2+\|g\|^2\right)\end{aligned}$$

da cui segue la tesi. □

Il fatto che tale identità rappresenti una condizione *necessaria* permette di individuare immediatamente immediatamente norme non indotte da prodotti scalari, verificando proprio che tali norme non rispettano tale proprietà.

Esempio 1.5. Mostriamo ad esempio due norme notevoli che **non** rispettano l'identità sopra dimostrata:

1. in $C^0([0,1])$, sia $\|f\|_U = \max_{x\in[0,1]} |f(x)|$ una norma. Si può far vedere che l'identità 1.1 non vale per tutte le funzioni dello spazio. Infatti, scelte $f(x) = 1-x$ e $g(x) = x$, si ha
$$\|f+g\|^2+\|f-g\|^2 = 1+1 \neq 2(1+1) = 2(\|f\|^2+\|g\|^2), \tag{1.7}$$
quindi tale norma non può essere indotta da alcun prodotto scalare.

2. in $\ell^p(\mathbb{N})$, si consideri la norma p-esima, definita come
$$\|a\|_p := \left(\sum_{i=1}^{\infty} |a_i|^p\right)^{1/p}. \tag{1.8}$$
Prese le successioni $a = (1,0,0,...)$ e $b = (0,1,0,...)$, si trova che l'identità 1.1 è soddisfatta solo nel caso $p=2$. Questo mostra come ℓ^p con $p \neq 2$ non sia uno spazio di Hilbert. Per dimostrare che esso sia uno spazio di Hilbert nel caso $p=2$ sono necessarie considerazioni più avanzate che vanno oltre i nostri attuali scopi.

Disuguaglianze notevoli

In questo paragrafo si vuole dimostrare che effettivamente la funzione introdotta nella Definizione 1.2 è una norma. Per far questo, resta soltanto da provare la disuguaglianza triangolare. In primis, cominciamo col dimostrare la **disuguaglianza di Cauchy-Schwarz**. [1]

Proposizione 1.1 (Cauchy-Schwarz). *Se $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$, ovvero la norma è indotta da un prodotto scalare, allora*
$$|\langle f,g\rangle| \leq \|f\|\cdot\|g\| \quad \forall f,g \in H. \tag{1.9}$$

Dimostrazione. Nel caso banale $\langle f,g\rangle = 0$, otteniamo $0 \leq 0$.
Per $\langle f,g\rangle \neq 0$, si può usare la positività del prodotto scalare per scrivere

$$\begin{aligned}0 \leq \langle f+\alpha g, f+\alpha g\rangle &= \langle f,f\rangle + \alpha\langle f,g\rangle + \overline{\alpha}\langle g,f\rangle + |\alpha|^2\langle g,g\rangle \\ &= \|f\|^2 + \overline{\alpha}\langle g,f\rangle + \alpha\langle f,g\rangle + |\alpha|^2\|g\|^2\end{aligned} \tag{1.10}$$

[1] *Hermann Schwarz*, da non confondere con il *Laurent Schwartz* delle distribuzioni e degli spazi $\mathscr{S}$.

valido $\forall \alpha \in \mathbb{C}$. In particolare, scelto $\alpha = -\|f\|^2/\langle f,g\rangle$ e semplificando, si ottiene la tesi. □

Si può ora dimostrare la **disuguaglianza triangolare**.

Proposizione 1.2 (Disuguaglianza triangolare). *Se $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$, allora*

$$\|f+g\| \leq \|f\| + \|g\| \quad \forall f,g \in H. \tag{1.11}$$

Dimostrazione. Si ha che

$$\begin{aligned} 0 \leq \|f+g\|^2 = \langle f+g, f+g\rangle = \|f\|^2 + \langle f,g\rangle + \langle g,f\rangle + \|g\|^2 \\ \leq \|f\|^2 + |\langle f,g\rangle| + |\langle g,f\rangle| + \|g\|^2 \\ \leq \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 \\ \leq (\|f\| + \|g\|)^2 \end{aligned} \tag{1.12}$$

ovvero la tesi. □

1.2 Base di uno spazio di Hilbert

Ricordiamo che una **base** di uno spazio vettoriale è un insieme di vettori che risultino al contempo linearmente indipendenti e un sistema di generatori dello spazio. Il numero di elementi in una base rappresenta la *dimensione* dello spazio.

Si introduce adesso il concetto di *base ortonormale.*

Definizione 1.3. *Se $\{v_i\}$ è una base dello spazio H, essa si dice* **ortonormale** *se $\langle v_i, v_j\rangle = \delta_{ij}$.*

Intuitivamente, si può pensare alla richiesta di ortonormalità dei vettori di base avendo in mente lo spazio euclideo $\mathbb{R}^n$. In quel caso, la proprietà che definisce una base ortonormale può essere interpretata come una richiesta di mutua ortogonalità tra coppie di vettori della base, da cui il nome. È semplice dimostrare la seguente

Proposizione 1.3. *Se un sistema di generatori è ortonormale, allora essi sono linearmente indipendenti.*

Dimostrazione. Dobbiamo semplicemente far vedere che l'unico modo per scrivere l'elemento nullo, tramite una combinazione lineare dei generatori, è quello in cui tutti i coefficienti sono nulli. Partendo dall'uguaglianza $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$, e facendone il prodotto scalare per v_i, si ha $\langle v_i, a_1 v_1 + \cdots + a_n v_n\rangle = \langle v_i, 0\rangle \Leftrightarrow a_i \langle v_i, v_i\rangle = 0 \Leftrightarrow a_i = 0 \ \forall i \in \{1, \ldots, n\}$, fatto che conferma la lineare indipendenza dei vettori $\{v_i\}$. □

In generale, il viceversa è chiaramente falso; in presenza di una base qualsiasi, tuttavia, è sempre possibile estrarre una base ortonormale costituita da vettori a norma unitaria tramite l'applicazione di alcuni "semplici" algoritmi. Il più noto di essi (ma non il più efficiente) prende il nome di **procedura di Gram-Schmidt**. Per ragioni di chiarezza, descriviamo la procedura nel caso di due vettori.

Proposizione 1.4 (Procedura di Gram-Schmidt). *Si abbia una base non ortonormale formata da due vettori v_1, v_2. Si applichi il seguente algoritmo:*

1. *si sceglie uno dei due vettori, ad esempio v_1, e lo si normalizza dividendo per la sua norma $\|v_1\|$. Si ottiene il versore $\hat{c}_1$;*

2. *si scompone v_2 nella componente parallela alla direzione di $\hat{c}_1$ e nella componente perpendicolare a essa,*

ovvero $v_2 = v_{2,\perp} + v_{2,//}$. Da questa, si sottrae la componente $v_{2,//}$ ottenendo $v_{2,\perp} = v_2 - \langle \hat{c}_1, v_2 \rangle \hat{c}_1$ e si normalizza, ottenendo il versore $\hat{c}_2$.

L'insieme dei due versori $\{\hat{c}_1, \hat{c}_2\}$ forma una base ortonormale.

Dimostrazione. Non è difficile mostrare che i due vettori così ottenuti sono ortogonali tra loro. Infatti

$$\begin{aligned} \|v_1\| \|v_{2,\perp}\| \langle \hat{c}_1, \hat{c}_2 \rangle &= \langle v_1, v_{2,\perp} \rangle = \langle v_1, v_2 - \langle \hat{c}_1, v_2 \rangle \hat{c}_1 \rangle \\ &= \langle v_1, v_2 \rangle - \langle \hat{c}_1, v_2 \rangle \langle v_1, \hat{c}_1 \rangle \\ &= \|v_1\| \langle \hat{c}_1, v_2 \rangle - \|v_1\| \langle \hat{c}_1, v_2 \rangle = 0 \end{aligned} \tag{1.13}$$

□

Data la natura *algoritmica* della procedura, è evidente che la stessa non può essere esplicitamente applicata nel caso di uno spazio di Hilbert infinito-dimensionale, caso in cui sarebbero necessari infiniti passi dell'algoritmo per convergere a una base ortonormale. In ogni caso, assumiamo che tramite la procedura di Gram-Schmidt possa essere sempre prodotta, anche se non in forma esplicita, una base ortonormale dello spazio di Hilbert, indipendentemente dalla sua dimensione. Nel seguito, lavoreremo sempre con basi ortonormali.

L'utilità dell'utilizzo di una base ortonormale consiste nel fatto di poter espandere ogni elemento f di uno spazio di Hilbert H come

$$\boxed{f = \sum_{i=1}^{\infty} c_i v_i, \quad c_i := \langle v_i, f \rangle .} \tag{1.14}$$

Questa espressione viene anche detta *espansione in serie formale di Fourier*. Il significato di questo nome sarà chiaro nel seguito.

1.2.1 Spazi di Hilbert a dimensione infinita

In spazi di Hilbert a dimensione finita, qualsiasi insieme di vettori linearmente indipendenti costituisce un sistema di generatori per lo spazio se tale insieme contiene un numero di elementi uguale alla dimensione dello spazio.

Nel caso di uno spazio di Hilbert a **dimensione infinita**, quest'ultima affermazione non è necessariamente vera. Specificamente, si può verificare che non tutti gli insiemi formati da infiniti vettori linearmente indipendenti sono generatori di uno spazio di Hilbert infinito-dimensionale.

Esempio 1.6. In $\ell^2(\mathbb{N})$, si può costruire un insieme di infiniti vettori linearmente indipendenti considerando tutte le successioni $\hat{e}_{2i}$, $i \geq 1$ tali da essere sempre uguali a zero tranne che in corrispondenza del $2i$-esimo elemento, che risulta uguale a 1. Non è difficile convincersi che la successione $(1, 0, 0, \dots)$ non può essere scritta in serie formale di Fourier utilizzando solo i vettori $\{\hat{e}_{2i}\}$. Tale insieme pertanto non costituisce una base.

È necessario dare qualche condizione in più sui vettori affinché si possa essere sicuri che costituiscano una base. Si cominci col dimostrare la seguente

Teorema 1.1 (Dis. di Bessel). *Sia $\mathcal{F} = \{\varphi_i\}_{i \in \mathbb{N}}$ è un insieme di vettori ortonormali di H. $\forall f \in H$ si ha*

$$\sum_{i=1}^{\infty} |\langle \varphi_i, f \rangle|^2 \leq \|f\|^2 \tag{1.15}$$

Dimostrazione. Definiamo $P_n := f - \sum_{i=1}^{n} \langle \varphi_i, f \rangle \varphi_i$. È ovvio che $P_n \in H$ in quanto combinazione lineare di elementi di H. Calcoliamo adesso la norma quadra di P_n, ovvero $\langle P_n, P_n \rangle$. Si ha che

$$\begin{aligned} 0 \leq \langle P_n, P_n \rangle &= \left\langle f - \sum_{i=1}^{n} \langle \varphi_i, f \rangle \varphi_i \, , \, f - \sum_{i=1}^{n} \langle \varphi_i, f \rangle \varphi_i \right\rangle = \\ &= \|f\|^2 - \sum_{j=1}^{n} \langle \varphi_j, f \rangle \langle f, \varphi_j \rangle - \cancel{\sum_{i=1}^{n} \langle f, \varphi_i \rangle \langle \varphi_i, f \rangle} + \cancel{\sum_{i,j=1}^{n} \langle f, \varphi_i \rangle \langle \varphi_j, f \rangle \langle \varphi_i, \varphi_j \rangle} \\ &= \|f\|^2 - \sum_{j=1}^{n} |\langle \varphi_j, f \rangle|^2 \xrightarrow[n,\infty]{} \|f\|^2 - \sum_{j=1}^{\infty} |\langle \varphi_j, f \rangle|^2 \end{aligned}$$

come volevasi dimostrare. □

Poiché i vettori in H hanno norma finita, la disuguaglianza di Bessel permette di affermare che la serie utilizzata nella dimostrazione precedente è di Cauchy. Questo fatto rende lecita la definizione della successione

$$\boxed{\begin{aligned} &P_n : H \longrightarrow H \\ &\qquad f \longmapsto \{P_n(f)\}_{n \in \mathbb{N}} := \sum_{k=1}^{n} \langle \varphi_k, f \rangle \varphi_k \end{aligned}} \tag{1.16}$$

Non è difficile far vedere che tale successione è di Cauchy. Infatti, se $n > m$,

$$\boxed{\|P_n - P_m\|^2 = \langle P_n - P_m, P_n - P_m \rangle = \sum_{k,l=m+1}^{n} \overline{\langle \varphi_k, f \rangle} \langle \varphi_l, f \rangle \langle \varphi_k, \varphi_l \rangle = \sum_{k=m+1}^{n} |\langle \varphi_k, f \rangle|^2 < \epsilon^2}$$

per ogni ϵ *piccolo a piacere* con m ed n *grandi abbastanza*, essendo la successione a quadrato sommabile. Dato che H è uno spazio metrico completo, tale successione risulta essere convergente a un qualche elemento $\tilde{f} \in H$. Non è detto, però, che tale $\tilde{f}$ coincida con f. Per garantire tale uguaglianza, bisogna introdurre un'ulteriore ipotesi sulla base ortonormale $\mathcal{F}$.

Definizione 1.4. *Dato uno spazio di Hilbert H, una base $\mathcal{F} = \{\varphi_i \in H : \langle \varphi_i, \varphi_j \rangle = \delta_{ij}\}$ si dice* **totale** *o* **completa** *se l'unico vettore di H contemporaneamente ortogonale a tutti i vettori della base è il vettore nullo, ovvero se $\forall i \in \mathbb{N}$*

$$\langle f, \varphi_i \rangle = 0 \iff f \equiv 0 \tag{1.17}$$

In questo caso, la base si indicherà come $\mathcal{F}_\varphi$.

Si può a questo punto dimostrare che la condizione di totalità è necessaria e sufficiente per garantire la convergenza della successione P_n sopra definita a f, ovvero $\hat{f} = f$.

Teorema 1.2. *Sia $\mathcal{F}_\varphi = \{\varphi_i \in H : \langle \varphi_i, \varphi_j \rangle = \delta_{ij}\}$. Le seguenti condizioni sono equivalenti.*

1. *$\mathcal{F}_\varphi$ è totale,*

2. $f = \sum_{i=1}^{\infty} \langle \varphi_i, f \rangle \varphi_i$,

3. $\langle f, g\rangle = \sum_{i=1}^{\infty} \langle f, \varphi_i\rangle \langle \varphi_i, g\rangle \; \forall f, g \in H$,

4. $\|f\|^2 = \sum_{i=1}^{\infty} |\langle f, \varphi_i\rangle|^2 \quad \forall f \in H$ (uguaglianza di Parseval).

Dimostrazione. In questi casi, la dimostrazione è *ciclica*.

$(1 \Rightarrow 2)$ Notiamo che $\left\langle \varphi_i, f - \hat{f} \right\rangle = \langle \varphi_i, f\rangle - \left\langle \varphi_i, \hat{f} \right\rangle = \langle \varphi_i, f\rangle - \sum_{j=1}^{+\infty} \langle \varphi_i, f\rangle \langle \varphi_i, \varphi_j\rangle = \langle \varphi_i, f\rangle - \langle \varphi_i, f\rangle = 0 \; \forall i \in \mathbb{N}$. Poiché per ipotesi l'unico vettore contemporaneamente ortogonale a tutti i vettori di base è il vettore nullo, si ha $f - \hat{f} = 0$ da cui $\hat{f} = f$.

$(2 \Rightarrow 3)$ $\langle f, g\rangle = \left\langle \sum_{i=1}^{+\infty} \langle \varphi_i, f\rangle \varphi_i, g \right\rangle = \sum_{i=1}^{+\infty} \overline{\langle \varphi_i, f\rangle} \langle \varphi_i, g\rangle = \sum_{i=1}^{+\infty} \langle f, \varphi_i\rangle \langle \varphi_i, g\rangle$.

$(3 \Rightarrow 4)$ Presa $f = g$, si trova l'uguaglianza richiesta.

$(4 \Rightarrow 1)$ Sia $f \in H : \langle f, \varphi_i\rangle = 0 \; \forall i \in \mathbb{N}$. Quindi, $\|f\| = 0 \iff f = 0$.

□

1.3 Prodotto scalare non-standard in uno spazio di Hilbert

Il prodotto scalare che può essere introdotto su uno spazio affinché esso risulti uno spazio di Hilbert non è unico. In generale, gli esempi di prodotto scalare dati in precedenza rappresentano, nei rispettivi contesti, delle scelte *naturali* di prodotto scalare. Descriviamo qui brevemente come è possibile modificare alcuni di tali prodotti al fine di estenderli in maniera non banale.

Consideriamo tre esempi.

Esempio 1.7. In $H = \mathbb{C}^n$, si definisca l'applicazione

$$\langle\langle a, b\rangle\rangle := \sum_{i=1}^{n} \overline{a_i} b_i c_i \tag{1.18}$$

dove c_i è una qualsiasi succesione non negativa, ovvero tale per cui $|c_i| > 0 \; \forall i$. Si dimostra che un'espressione di questo tipo definisce ancora un prodotto scalare, rendendo in particolare tale insieme uno spazio di Hilbert. È chiaro che, se poniamo $c_i \equiv 1 \; \forall i \in \mathbb{N}$, otteniamo il prodotto scalare standard precedentemente introdotto.

Esempio 1.8. L'applicazione (1.18), con $n = \infty$, può essere usata come prodotto scalare nello spazio $\ell^2(\mathbb{N})$, rendendolo uno spazio di Hilbert. In questo caso, la disuguaglianza di Cauchy-Schwarz diventa

$$\left| \sum_{i=1}^{\infty} \overline{a_i} b_i c_i \right| \leq \left(\sum_{i=1}^{\infty} |a_i|^2 c_i \right) \left(\sum_{i=1}^{\infty} |b_i|^2 c_i \right) \tag{1.19}$$

In generale, la successione c_i ha a tutti gli effetti il ruolo di un **peso**, che dona ad ognuno dei termini $\overline{a_i} b_i$ un'importanza diversa nel calcolo del valore finale.

Esempio 1.9. Seguendo quanto fatto finora con l'introduzione di un *peso* nel prodotto scalare, ridefiniamo il prodotto scalare in $\mathscr{L}^2(D)$ come

$$\langle\langle f, g\rangle\rangle = \int_D \overline{f(x)} g(x) p(x) \, \mathrm{d}x \,, \tag{1.20}$$

con $p(x) > 0$ q.o. in D. L'insieme delle funzioni a quadrato sommabile dotato di questo prodotto scalare continua a rappresentare uno spazio di Hilbert.

1.4 Polinomi ortogonali

L'idea dietro alla costruzione di una famiglia di **polinomi ortogonali** risiede nel fatto che, in opportuni spazi di Hilbert, è possibile costruire una famiglia di vettori ortogonali fra loro usando funzioni polinomiali di ordine via via crescente. La loro introduzione semplifica molte applicazioni.

Solitamente, i polinomi ortogonali vengono costruiti nello spazio

$$\mathscr{L}^2_w(a,b) = \left\{ f(x) \;\; \mathscr{L}\text{-misurabile} : \|f\|^2_w := \int_a^b |f(x)|^2 w(x) \, \mathrm{d}x < +\infty \right\} \tag{1.21}$$

con $-\infty \le a \le b \le +\infty$. Qui, la funzione $w(x)$ è un'opportuna *funzione peso* che per consistenza deve soddisfare le condizioni

$$w(x) > 0 \quad \text{e} \quad \int_a^b x^n w(x) \, \mathrm{d}x < \infty \quad \forall n \in \mathbb{N} \cup \{0\}, \tag{1.22}$$

che equivalgono a richiedere che la funzione peso sia una funzione positiva con tutti i momenti finiti. Notiamo che se $w(x) \in \mathscr{L}^1$, la finitezza dei momenti è sempre verificata dato che

$$\left| \int_a^b x^n w(x) \mathrm{d}x \right| \le |b-a|^n \|w\| < \infty. \tag{1.23}$$

Nello spazio $\mathscr{L}^2_w(a,b)$, il prodotto scalare è definito sotto forma di un prodotto scalare non-standard del tipo

$$\langle f, g\rangle_w = \int_a^b \overline{f(x)} g(x) \, w(x) \, \mathrm{d}x \equiv \left\langle \sqrt{w(x)} f(x), \sqrt{w(x)} g(x) \right\rangle, \tag{1.24}$$

dove $\langle \cdot, \cdot \rangle$ rappresenta il prodotto scalare standard in $\mathscr{L}^2(a,b)$.

Si può ricavare un set di polinomi ortogonali entro questo spazio applicando la procedura di ortogonalizzazione di Gram-Schmidt all'insieme $\{1, x, x^2, \dots\}$. Partendo per esempio dai primi due termini:

$$p_0(x) = \frac{1}{\|1\|_w}; \; P_1(x) = x - \langle p_0, x\rangle_w \, p_0(x) \Rightarrow p_1(x) = \frac{P_1(x)}{\|P_1(x)\|_w}; \; \cdots \tag{1.25}$$

Il risultato di tale procedura dipende fortemente dalla scelta della funzione peso $w(x)$. Si osserva però che, indipendentemente da essa, si ha che

1. $p_n(x)$ è un polinomio di grado n,
2. il coefficiente della potenza n-esima in $p_n(x)$ è uguale a $\frac{1}{\|P_n(x)\|_w}$.

Noti che siano i coefficienti di ogni polinomio, essi possono essere rappresentati in forma matriciale come

$$\begin{pmatrix} p_0(x) \\ p_1(x) \\ p_2(x) \\ \vdots \\ p_n(x) \end{pmatrix} = \begin{pmatrix} \|P_0(x)\|_w^{-1} & 0 & 0 & \dots & 0 \\ a_{10} & \|P_1(x)\|_w^{-1} & 0 & \dots & 0 \\ a_{20} & a_{21} & \|P_2(x)\|_w^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n0} & a_{n1} & a_{n2} & \dots & \|P_n(x)\|_w^{-1} \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$$

ovvero

$$\mathbf{P} = M\mathbf{X}. \tag{1.26}$$

È immediato notare che, poiché tutti i termini sulla diagonale di M sono positivi, la matrice ha determinante non nullo e quindi risulta invertibile. Questo permette di esprimere le potenze di x come combinazioni lineare di polinomi ortogonali, ottenendo $\mathbf{X} = M^{-1}\mathbf{P}$.

Dimostriamo ora la seguente proposizione.

Proposizione 1.5. *Sia $q_l(x)$ un qualsiasi polinomio di grado $l < n$. Allora $\langle q_l, p_n \rangle_w = 0$.*

Dimostrazione. Indicando con q_k i coefficienti del polinomio $q_l(x)$, si può scrivere

$$\langle q_l, p_n \rangle_w = \sum_{k=0}^{l} q_k \left\langle x^k, p_n \right\rangle_w .$$

In virtù della considerazione precedente sull'invertibilità della matrice M, è possibile espandere il termine $\langle x^k, p_n \rangle_w$ come

$$\left\langle x^k, p_n \right\rangle_w = \sum_{j=0}^{k} c_j \langle p_j, p_n \rangle_w = \sum_{j=0}^{k} c_j \delta_{j,n} = 0$$

dato che $k < n$. □

Ogni famiglia di polinomi ortogonali è definita a meno di una costante moltiplicativa, come riassunto nella seguente

Proposizione 1.6. *Si consideri un insieme di polinomi $\{Q_n(x) : \langle Q_n, Q_m \rangle = 0 \text{ se } n \neq m\}$ che soddisfa la proprietà che il grado di $Q_n(x)$ sia n. Allora i polinomi della famiglia sono definiti a meno di una costante moltiplicativa.*

Una maniera rapida per calcolare esplicitamente i coefficienti del polinomio n-esimo di una famiglia ortogonale è rappresentata dalla seguente **formula di ricorrenza a tre termini**:

Proposizione 1.7. *Se $\{p_n\}$ è una famiglia di polinomi ortogonali, allora vale che*

$$x\, p_n(x) = a_n\, p_{n+1}(x) + b_n\, p_n(x) + c_n\, p_{n-1}(x) \tag{1.27}$$

Dimostrazione. Si consideri $x\, p_n(x) - a_n p_{n+1}(x) = \sum_{l=0}^{n} \alpha_l\, p_l(x)$. Si può scegliere il coefficiente a_n in maniera tale che il primo membro dell'equazione sia un polinomio di grado n. Moltiplicando

scalarmente per un generico $p_j(x)$, si ha

$$\left\langle p_j, x\, p_n(x) - a_n p_{n+1}(x) \right\rangle_w = \sum_{l=0}^{n} \alpha_l \left\langle p_j, p_l \right\rangle_w = \begin{cases} 0 & j \le n-2 \\ \neq 0 & \text{altrimenti} \end{cases},$$

a causa del fatto che $\left\langle p_j, x\, p_n \right\rangle_w - a_n \cancel{\left\langle p_j, p_{n+1} \right\rangle_w} = \left\langle x\, p_j, p_n \right\rangle_w = 0$ per $j \le n-2$, come conseguenza della proposizione precedente. Quindi in generale si ha $x\, p_n(x) - a_n\, p_{n+1}(x) = \alpha_{n-1}\, p_{n-1}(x) + \alpha_n\, p_n(x)$, ovvero la tesi. □

1.4.1 Formula di Rodrigues

Le famiglie di *polinomi ortogonali classici*, cui ci limiteremo nel seguito, possono essere ottenute a partire da una formula del tutto generale, detta **formula di Rodrigues**[2].

Proposizione 1.8 (Formula di Rodrigues). *Ogni termine di una famiglia di* polinomi ortogonali classici *può essere espressa come*

$$P_n(x) = \frac{1}{w(x)} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left[w(x)\, Q^n(x) \right] \tag{1.28}$$

dove $w(x)$ è la funzione peso che definisce lo spazio di lavoro $\mathscr{L}^2_w(a,b)$ e $Q(x)$ è un polinomio a valori non negativi nell'intervallo $[a,b]$.

Si può dimostrare, anche con l'ausilio dell'induzione, la seguente

Proposizione 1.9. *Scelti $Q(x), w(x)$ polinomi tali che* $\deg P_1 = 1$*, ovvero il grado del primo polinomio sia uguale a 1, si ha sempre* $\deg P_n = n$*.*

Esempio 1.10. Si mostri per induzione che, ad esempio scegliendo $Q(x) = 1$ e $w(x) = \mathrm{e}^{-x^2}$, si ottiene che P_n è un polinomio di grado n. La formula (1.8) diventa

$$P_n(x) = \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2} \tag{1.29}$$

È banale verificare che $P_1(x)$ sia di grado 1. Supponiamo che P_{n-1} sia di grado $n-1$ e mostriamo che P_n è di grado n. Da (1.29) si ha $\mathrm{e}^{-x^2} P_{n-1}(x) = \frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}} \left(\mathrm{e}^{-x^2} \right)$. Allora:

$$\begin{aligned} P_n(x) &= \mathrm{e}^{x^2} \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}} \mathrm{e}^{-x^2} \right) = \mathrm{e}^{x^2} \frac{\mathrm{d}}{\mathrm{d}x} \left(\mathrm{e}^{-x^2} P_{n-1}(x) \right) \\ &= \mathrm{e}^{x^2} \left(-2x\, \mathrm{e}^{-x^2} P_{n-1}(x) + \mathrm{e}^{-x^2} \frac{\mathrm{d}}{\mathrm{d}x} P_{n-1}(x) \right) \\ &= -2x\, P_{n-1}(x) + \frac{\mathrm{d}}{\mathrm{d}x} P_{n-1}(x) \end{aligned}$$

che è chiaramente un polinomio di grado n.

Solitamente, il polinomio $Q(x)$ è al massimo del secondo ordine. In questo caso, le possibili scelte di $Q(x)$ sono

[2] *Benjamin Olinde Rodrigues*, banchiere, matematico e riformista sociale francese, ★1795 - †1851. È noto, misteriosamente, anche come *Rodriguez* (pronunciato come in *Belén Rodríguez*, ndr.)

1. $Q(x) = 1$,

2. $Q(x) = x - a$,

3. $Q(x) = (b - x)(x - a)$.

Si ricavino adesso i pesi corrispondenti a queste tre scelte di $Q(x)$. Nell'ipotesi che $P_1(x)$ sia un polinomio di primo grado, la formula di Rodrigues restituisce un'equazione differenziale del primo ordine in $w(x)$:

$$\begin{aligned} A\,x + B &= \frac{1}{w(x)}\frac{\mathrm{d}}{\mathrm{d}x}\Big[w(x)\,Q(x)\Big] = \frac{1}{w(x)}\Big[w'(x)\,Q(x) + w(x)\,Q'(x)\Big] \\ &= \frac{w'(x)\,Q(x)}{w(x)} + Q'(x) \end{aligned}$$

da cui

$$\frac{w'(x)}{w(x)} = \frac{A\,x + B - Q'(x)}{Q(x)}, \tag{1.30}$$

che può essere risolta per ricavare la funzione peso $w(x)$. Ad ogni scelta di $Q(x)$ corrisponde anche un diverso dominio di integrazione, che può essere a misura finita o infinita, e delle condizioni, dettate dalla richiesta di convergenza dell'integrale, sui parametri arbitrari derivanti dalla risoluzione dell'equazione differenziale (1.30). In particolare, troviamo che

1. da $Q(x) = 1$, si ottiene $w(x) = \exp\{A\,\frac{x^2}{2} + B\,x + C\}$, con $A < 0$, $a = -\infty$, $b = +\infty$.

2. da $Q(x) = x - a$, si ottiene $w(x) = \exp\{A\,x\}(x - a)^B$, con $A < 0 \wedge B > -1$ e $|a| < +\infty, b = +\infty$.

3. da $Q(x) = (b - x)(x - a)$, si ottiene $w(x) = (x - a)^B(b - x)^A$, con $A, B > -1$, $b - a < \infty$.

Ricordiamo che a e b sono gli estremi dell'intervallo su cui sono definite le funzioni nello spazio di lavoro $\mathscr{L}^2_w(a, b)$. Le condizioni riportate sono necessarie per garantire la convergenza dell'integrale di $w(x)$.

Si dimostra che un $P_n(x)$ così ricavato è ortogonale a qualsiasi polinomio di grado inferiore. Quindi, polinomi prodotti dalla formula (1.8) formano effettivamente famiglie di polinomi ortogonali. Considerando che polinomi ortogonali della stessa famiglia sono definiti a meno di una costante additiva, si può riscrivere la (1.8) come

$$p_n(x) = \frac{1}{k_n}\frac{1}{w(x)}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[w(x)\,Q^n(x)\right] \tag{1.31}$$

in maniera da potere fissare arbitrariamente la normalizzazione. Manipolando quest'ultima espressione, si può mostrare che i polinomi ortogonali vengono fuori come soluzioni di equazioni differenziali del tipo

$$\begin{aligned} &Q(x)p''_n(x) + k_{1'}\,p_1(x)p'_n(x) + \lambda_n p_n(x) = 0, \quad \forall n \geq 2 \\ &\text{con} \quad \lambda_n = -n\left(k_1 p'_1(x) + Q''(x)\frac{n-1}{2}\right). \end{aligned} \tag{1.32}$$

Tali equazioni rappresentano un'altra maniera di definire una famiglia di polinomi ortogonali.

1.4.2 Famiglie notevoli di polinomi ortogonali

In questa sezione si considerano alcuni esempi di *polinomi ortogonali classici*, ovvero ottenibili tramite la formula di Rodrigues. Alcuni di essi sono di particolare interesse nelle applicazioni.

1. **Polinomi di Legendre**: scelti $a = -1$ e $b = 1$, si ottiene $Q(x) = 1 - x^2$. Inoltre, ponendo $A = B = 0$, si ottiene $w(x) = 1$. La formula (1.31) si scrive allora come

$$p_n(x) = \frac{1}{k_n} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left[(1 - x^2)^n \right], \tag{1.33}$$

dove k_n può essere scelto ad arbitrio, ma viene solitamente scelto in maniera da soddisfare la condizione $p_n(1) = 1$. Si ottengono così i **polinomi di Legendre**:

$$\begin{array}{ccc} p_0(x) = 1 & p_1(x) = x & p_2(x) = \frac{1}{2}(3x^2 - 1) \\ p_3(x) = \frac{1}{2}(5x^3 - x) & p_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3) & p_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x), \dots \end{array}$$

Si noti che hanno parità definita, ovvero vale che $p_n(-x) = (-1)^n\, p_n(x)$. Si riporta anche la formula di ricorrenza e l'equazione differenziale da cui sono ottenibili:

$$(n+1)p_{n+1}(x) = (2n+1)xp_n(x) - np_{n-1}(x), \tag{1.34}$$

$$(1 - x^2)p_n'' - 2xp_n' + n(n+1)p_n = 0. \tag{1.35}$$

I polinomi di Legendre sono impiegati ad esempio nello sviluppo in multipoli del potenziale elettrostatico.

2. **Polinomi di Chebyshev (prima specie)**: scelti $a = -1$ e $b = 1$ e ponendo $A = B = -1/2$, si ottiene $w(x) = (1 - x^2)^{1/2}$. La formula (1.31) si scrive allora come

$$T_n(x) = \frac{1}{k_n} \sqrt{1 - x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left[(1 - x^2)^{n - 1/2} \right] \tag{1.36}$$

k_n viene scelto in maniera da soddisfare la condizione $T_n(1) = 1$. Si ottengono i **polinomi di Chebyshev della prima specie**:

$$\begin{array}{ccc} T_0(x) = 1 & T_1(x) = x & T_2(x) = 2x^2 - 1 \\ T_3(x) = 4x^3 - x & T_4(x) = 8x^4 - 8x^2 + 1 & T_5(x) = 16x^5 - 20x^3 + 5x, \dots \end{array}$$

Si nota che hanno parità definita, ovvero vale che $T_n(-x) = (-1)^n T_n(x)$. Si riporta anche la formula di ricorrenza e l'equazione differenziale da cui sono ottenibili:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \tag{1.37}$$

$$(1 - x^2)T_n'' - xT_n' + n^2 T_n = 0. \tag{1.38}$$

3. **Polinomi di Chebyshev (seconda specie)**: scelti $a = -1$ e $b = 1$ e ponendo $A = B = 1/2$, si ottiene $w(x) = (1 - x^2)^{1/2}$. La formula (1.31) si scrive allora come

$$U_n(x) = \frac{1}{k_n} \frac{1}{\sqrt{1 - x^2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left[(1 - x^2)^{n + 1/2} \right] \tag{1.39}$$

k_n viene solitamente scelto in maniera da soddisfare la condizione $U_n(1) = 1$. Si ottengono i **polinomi di Chebyshev di seconda specie**:

$$U_0(x) = 1 \qquad U_1(x) = x \qquad U_2(x) = \tfrac{1}{3}(4x^2 - 1)$$
$$U_3(x) = 2x^3 - x \quad U_4(x) = \tfrac{1}{5}(16x^4 - 12x^2 + 3) \quad \dots$$

Si nota che hanno parità definita, ovvero vale che $U_n(-x) = (-1)^n U_n(x)$. Si trascrive anche la formula di ricorrenza e l'equazione differenziale da cui sono ottenibili:

$$(n+2)U_{n+1}(x) = 2x(n+1)U_n(x) - nU_{n-1}(x), \tag{1.40}$$
$$(1-x^2)U_n'' - 3xU_n' + n(n+2)U_n = 0. \tag{1.41}$$

4. **Polinomi di Laguerre generalizzati**: scelti $a = 0$ e $b = +\infty$, si ottiene $w(x) = e^{-x}x^\alpha$. La formula (1.31) si scrive allora come

$$L_n^\alpha(x) = \frac{1}{n!}e^x x^{-\alpha}\frac{d^n}{dx^n}\left[e^{-x}\,x^{\alpha+n}\right] \tag{1.42}$$

avendo posto $k_n = n!$ come normalizzazione. Si ottengono i **polinomi di Laguerre generalizzati**:

$$L_0^\alpha(x) = 1, \quad L_1^\alpha(x) = 1 - (\alpha+10)x, \quad L_2^\alpha(x) = \frac{1}{2}x^2 - (2+\alpha)x + \frac{1}{2}(\alpha^2 + 3\alpha + 2)$$

e così via. Scriviamo anche la formula di ricorrenza e l'equazione differenziale da cui sono ottenibili:

$$(n+1)L_{n+1}^\alpha(x) = (2n + \alpha - 1 + x)L_n^\alpha(x) - (n+\alpha)L_{n-1}^\alpha(x), \tag{1.43}$$
$$xL_n^{\alpha\prime\prime} + (1 + \alpha - x)L_n^{\alpha\prime} + nL_n^\alpha = 0. \tag{1.44}$$

Tali polinomi sono usati nella scrittura della funzione d'onda elettronica dell'atomo di idrogeno.

5. **Polinomi di Hermite**: scelti $a = -\infty$ e $b = +\infty$ si può scegliere $Q(x) = 1$. Inoltre, ponendo $A = -2$ e $B = C = 0$, si ottiene $w(x) = e^{-x^2}$. La formula (1.31) si scrive allora come

$$H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2} \tag{1.45}$$

avendo scelto $k_n = (-1)^n$. Si ottengono i **polinomi di Hermite**[3]:

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 2(2x^2 - 1), \quad H_3(x) = 4(2x^3 - 3x), \quad \dots$$

Scriviamo anche la formula di ricorrenza e l'equazione differenziale da cui sono ottenibili:

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x), \tag{1.46}$$
$$H_n'' - 2xH_n' + 2nH_n = 0. \tag{1.47}$$

Questi polinomi hanno parità definitiva, in quanto

$$H_n(-x) = (-1)^n H_n(x) \tag{1.48}$$

Questi polinomi sono usati in meccanica quantistica nella scrittura delle funzioni d'onda dell'oscillatore armonico.

[3] I polinomi qui riporta sono i cosiddetti *polinomi di Hermite fisici*.

Capitolo 2

Serie di Fourier

L'argomento, che è possibile affrontare con una descrizione che prescinde dagli spazi di Hilbert, viene qui ripreso *molto* brevemente sotto un altro punto di vista che rende la trattazione più *semplice*.

2.1 Introduzione

L'ambiente di lavoro è $\mathscr{L}^2(-\pi,\pi)$, o, analogamente, $\mathscr{L}^2(k\pi,(k+1)\pi); k \in \mathbb{Z}$. Ci interesserà particolarmente lavorare con intervalli centrati nell'origine per proprietà che saranno chiare nel seguito.

Proposizione 2.1. *In* $\mathscr{L}^2(-\pi,\pi)$ *l'insieme*

$$\mathcal{F} := \left\{ \varphi_0 = \frac{1}{\sqrt{2\pi}}, \varphi_1 = \frac{1}{\sqrt{\pi}}\sin(x), \varphi_2 = \frac{1}{\sqrt{\pi}}\cos(x), \varphi_3 = \frac{1}{\sqrt{\pi}}\sin(2x), \dots \right\}$$

è una base ortonormale e totale rispetto al prodotto scalare standard.

L'ortonormalità può essere dimostrata semplicemente notando che, $\forall l, m \in \mathbb{N}$,

1. $\displaystyle\int_{-\pi}^{\pi} \sin(lx)\cos(mx)\mathrm{d}x = 0$

2. $\displaystyle\int_{-\pi}^{\pi} \sin(lx)\sin(mx)\mathrm{d}x = 2\pi\delta_{l,m}$

3. $\displaystyle\int_{-\pi}^{\pi} \cos(lx)\cos(mx)\mathrm{d}x = 2\pi\delta_{l,m}$

Si assume la completezza della base, fatto dimostrabile.

Ricordando il teorema (1.2), si può affermare che $\forall f \in \mathscr{L}^2(-\pi,\pi)$ esiste un'espansione formale $f = \sum_{i=0}^{+\infty} \langle \varphi_i, f \rangle \varphi_i$, dove $\varphi_i \in \mathcal{F}$. Scrivendo esplicitamente alcuni termini di questo sviluppo,

si ottiene

$$\begin{aligned} f(x) = &\left(\frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} f(x)\mathrm{d}x\right)\frac{1}{\sqrt{2\pi}} + \left(\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\sin(x)\mathrm{d}x\right)\frac{\sin(x)}{\sqrt{\pi}} \\ &+ \left(\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\cos(x)\mathrm{d}x\right)\frac{\cos(x)}{\sqrt{\pi}} + \left(\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\sin(2x)\mathrm{d}x\right)\frac{\sin(2x)}{\sqrt{\pi}} + \dots \end{aligned} \tag{2.1}$$

Si possono allora definire dei coefficienti

$$D_l := \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(lx)\mathrm{d}x \quad \text{e} \quad P_l := \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(lx)\mathrm{d}x \tag{2.2}$$

per cui la (2.1) diventa

$$\boxed{f(x) = \frac{1}{2}P_0 + \sum_{l=1}^{\infty}\left[D_l\sin(lx) + P_l\cos(lx)\right]} \tag{2.3}$$

che è quello che viene detto **sviluppo in serie di Fourier** di f. È evidente dalla definizione che se f è pari, allora $D_l = 0\ \forall\, l$. Al contrario, se f è dispari, $P_l = 0\ \forall\, l$.

2.2 Norma della funzione

Facendo riferimento a (2.3), si può esprimere la norma di f in relazione al suo sviluppo in serie di Fourier:

$$\begin{aligned} \|f\|^2 &= |\langle\varphi_0, f\rangle|^2 + |\langle\varphi_1, f\rangle|^2 + |\langle\varphi_3, f\rangle|^2 + \dots \\ &= \left|\frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} f(x)\mathrm{d}x\right|^2 + \left|\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\sin(x)\mathrm{d}x\right|^2 + \left|\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\cos(x)\mathrm{d}x\right|^2 + \dots \\ &= \left|\sqrt{\frac{\pi}{2}}P_0\right|^2 + |\sqrt{\pi}D_1|^2 + |\sqrt{\pi}P_1|^2 + \dots \end{aligned} \tag{2.4}$$

da cui la **formula di Parseval**

$$\boxed{\|f\|^2 = \pi\left[\frac{|P_0|^2}{2} + \sum_{l=1}^{+\infty}\left(|P_l|^2 + |D_l|^2\right)\right].} \tag{2.5}$$

Poiché $f \in \mathscr{L}^2(-\pi,\pi)$, allora le serie dei $|P_l|^2$ e dei $|D_l|^2$ sono convergenti. Questo implica che $\{D_l\}, \{P_l\} \in \ell^2(\mathbb{N})$. Questo fatto può essere utile per individuare immediatamente quelle serie trigonometriche che non possono essere sviluppi in serie di Fourier di alcuna funzione. Ad esempio, la serie $\sum_l \sin(lx) l^{\ 3/7}$ non può essere una serie di Fourier dato che $\sum_l l^{\ 6/7}$ non è convergente in quanto serie armonica generalizzata con esponente minore di 1. Il fatto che $\{D_l\}$ e $\{P_l\}$ siano quadrato sommabili implica inoltre che

$$\lim_{n,\infty}\int_{-\pi}^{\pi} f(x)\cos(nx)\mathrm{d}x = \int_{-\pi}^{\pi} f(x)\sin(nx)\mathrm{d}x = 0. \tag{2.6}$$

Esempio 2.1 (Funzione segno). Si sviluppi in serie di Fourier la funzione $f(x) = \text{sgn}(x) = \pm 1$ se $x \gtrless 0$ ristretta all'intervallo $[-\pi, \pi]$. Poiché f è dispari, sicuramente $P_n = 0$. Dal calcolo di D_n si ottiene invece

$$D_n = \begin{cases} 0 & \text{n pari,} \\ \frac{4}{\pi n} & \text{n dispari.} \end{cases} \tag{2.7}$$

Allora

$$\text{sgn(x)} = \frac{4}{\pi} \sum_{n=0}^{+\infty} \frac{1}{2n+1} \sin(2n+1)x. \tag{2.8}$$

Notiamo che per $x = 0$, la serie converge a 0, che è esattamente la media aritmetica dei limiti destro e sinistro della funzione in $x = 0$. Questo è un risultato generale, ovvero una serie di Fourier converge sempre puntualmente a tale valore, che coincide col valore vero della funzione soltanto nei punti in cui essa è continua.
Dall'uguaglianza di Parseval, si ha inoltre

$$\|f\|^2 = \frac{16}{\pi} \sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2}. \tag{2.9}$$

Ma $\|f\|^2 = 2\pi$, facilmente calcolabile da definizione di norma. Allora si ottiene

$$\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}. \tag{2.10}$$

Lo sviluppo in serie di Fourier può quindi a volte consentire il calcolo di particolari serie convergenti.
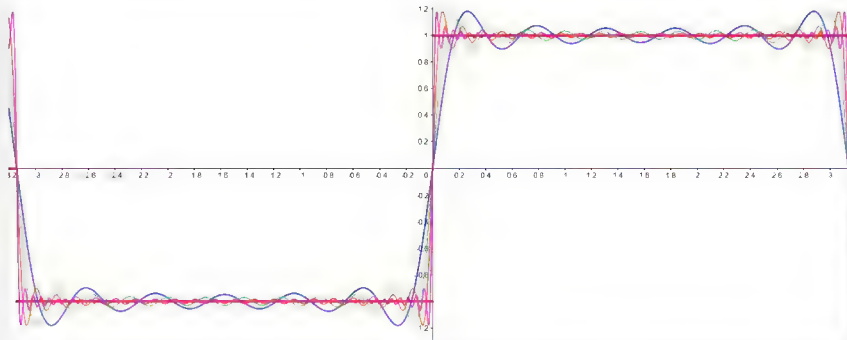
Figura 2.1: Funzione sgn(x) in $[-\pi, \pi]$ sovrapposta ai polinomi di Fourier di ordine 5, 10, 20 e 50. Si nota che in corrispondenza delle discontinuità di tipo salto, i polinomi di Fourier mostrano delle instabilità, dette **instabilità di Gibbs**.

Capitolo 3

Teoria degli operatori limitati

Le teoria degli operatori riveste un ruolo di fondamentale importanza nella maggior parte delle applicazioni fisico-matematiche. Nel corso del seguente capitolo vengono introdotti alcuni aspetti fondamentali, specializzando in seguito la trattazione agli operatori definiti su spazi di Hilbert.

3.1 Introduzione

Cominciamo dalle definizioni basilari. Il concetto di operatore è estremamente generale.

Definizione 3.1. *Un* **operatore** *è un oggetto T che mappa uno spazio vettoriale lineare (SVL) in un altro SVL. Il suo* **dominio** *è l'insieme* $\mathrm{dom}(T) = \{v \in V : \exists!\, Tv\}$.

Non è affatto detto che $\mathrm{dom}(T) = V$, cioè che il dominio coincida con lo spazio vettoriale di partenza.

Esempio 3.1. Se $V = C^1(\mathbb{R})$ e $T = \frac{\mathrm{d}}{\mathrm{d}x}$, l'operatore lineare T ha per dominio tutto V. Se invece consideriamo lo stesso operatore ma con $V = C^0(\mathbb{R})$, il dominio è contenuto in V in quanto esistono funzioni continue ma non derivabili.

Esempio 3.2. L'applicazione $N : V \to [0, +\infty[$ definita come $Nf = \langle f, f \rangle$ è un operatore. Esso associa ad ogni vettore dello spazio normato V la sua norma $||\cdot||$.

Tra tutti gli operatori, un'importante classe è costituita dagli operatori lineari.

Definizione 3.2. *Un* ***operatore*** *T viene detto* **lineare** *se è valida la proprietà*

$$T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2) \quad \forall\, \alpha_1, \alpha_2 \in \mathbb{C} \ \ e \ \ \forall\, v_1, v_2 \in V.$$

Esempio 3.3. Due operatori lineari "banali" ma fondamentali sono:

1. **operatore nullo**: $\widehat{0}$ tale che $\widehat{0}v = 0_V \quad \forall v \in V$;
2. **operatore identità**: $\mathbb{1}$ tale che $\mathbb{1}v = v \quad \forall v \in V$

Esempio 3.4. Sia $V = H$ uno spazio di Hilbert. Sia $T_{f_0} : H \to \mathbb{C}, f_0 \neq 0$ l'operatore definito come $T_{f_0} g = \langle f_0, g \rangle$. Esso risulta essere un operatore lineare, date le proprietà del prodotto scalare. Analogamente, l'operatore $V_{f_0} g = \langle g, f_0 \rangle$ risulterà antilineare, ovvero $V_{f_0}(\alpha_1 v_1 + \alpha_2 v_2) = \bar{\alpha}_1 V_{f_0}(v_1) + \bar{\alpha}_2 V_{f_0}(v_2)$.

3.2 Operatori su spazi di Hilbert

Particolarmente importanti sono gli operatori definiti da uno spazio di Hilbert in se stesso, oppure tra spazi di Hilbert diversi. Infatti, un operatore definito da H in H produce, come risultato, un vettore in H, su cui l'operatore può essere fatto agire un'altra volta. In altre parole, per gli operatori $T : H \to H$ è possibile definire le potenze di operatori e, tramite esse e gli sviluppi in serie di Taylor, le funzioni $f(T)$ di operatori[1]. Ammettendo che tale serie esista e converga, si pone la seguente definizione.

Definizione 3.3 (Funzione di un operatore). *Data una funzione f analitica in 0 e dato un operatore $T : H \to H$. La* **funzione dell'operatore** *$f(T)$ è definita come lo sviluppo in serie di Maclaurin*

$$f(T) := \sum_{i=0}^{\infty} \frac{f^{(n)}(0)}{n!} T^n. \tag{3.1}$$

Esempio 3.5. Si consideri l'operatore T_{f_0} introdotto in precedenza, modificato affinché mappi H in H

$$T_{f_0} g = \langle f_0, g \rangle f_0. \tag{3.2}$$

Un calcolo diretto mostra che

$$T_{f_0}^2 g = T_{f_0}(T_{f_0} g) = T_{f_0}(\langle f_0, g \rangle f_0) = \langle f_0, g \rangle \langle f_0, f_0 \rangle f_0 \tag{3.3}$$

che, se f_0 è un versore, diventa $T_{f_0}^2 g = T_{f_0} g$.

Si dà la seguente naturale definizione.

Definizione 3.4. *Due operatori sono uguali se, e solo se, applicati a un qualsiasi vettore del comune dominio, producono il medesimo risultato.*

Nell'esempio precedente, si può quindi concludere che $T_{f_0}^2 = T_{f_0}$. Questo fatto conduce ad un'altra importante definizione.

Definizione 3.5. *Un operatore A tale che $A^2 = A$ è detto* **idempotente** *o anche, nel contesto della meccanica quantistica,* **operatore di proiezione**.

Un operatore idempotente è, inoltre, tale che $A^n = A \ \forall n \in \mathbb{N}$, fatto facilmente dimostrabile per induzione.

3.2.1 Continuità di un operatore

Se un operatore è definito su uno spazio normato e a valori in uno spazio normato, allora è possibile introdurre il concetto di continuità dell'operatore, analogo a quello per le funzioni.

Definizione 3.6. *Si consideri $T : V \to W$, con V, W spazi normati. L'operatore T si dice* continuo *se $\forall \epsilon > 0 \ \exists \delta_{(\epsilon)} > 0 : \|x - x_0\|_V \le \delta_{(\epsilon)} \Rightarrow \|T(x) - T(x_0)\|_W \le \epsilon$.*

Similmente, un altro concetto che si può introdurre è quello di limitatezza dell'operatore.

[1] Il fatto che l'operatore sia definito da $H \to H$ non basta per assicurare che le sue potenze siano ben definite, ma è sicuramente una condizione necessaria affinché questo sia possibile.

Definizione 3.7. *L'operatore T viene detto* limitato *se* $\exists M > 0 : \forall x \in V \Rightarrow \|Tx\|_W \leq M\|x\|_V$.

Si può dimostrare la totale equivalenza tra le due definizioni. Per semplicità dimostriamo tale equivalenza solo nel caso di operatori lineari.

Proposizione 3.1. *Sia T un operatore. T è limitato $\iff$ T è continuo.*

Dimostrazione. $(\Rightarrow)$ Verifichiamo soltanto l'implicazione diretta. Consideriamo due elementi $x, y \in V$ tali che $\|x - y\|_W < \delta_\epsilon$ con $\delta_\epsilon > 0$. Se T è limitato con costante M, allora

$$\|T(x) - T(y)\|_W = \|T(x - y)\|_W \leq M\|x - y\|_W < \delta_\epsilon M < \epsilon \tag{3.4}$$

se δ_ϵ è scelta più piccola del valore ϵ/M. La $(\Leftarrow)$ non viene qui dimostrata. □

3.2.2 Norma di un operatore

Analogamente al concetto di continuità, si può introdurre la norma di un operatore. Per prima cosa, introduciamo lo spazio in cui vivono gli operatori lineari limitati definiti tra spazi normati.

Definizione 3.8. *Siano $T[\|\cdot\|_t], V[\|\cdot\|_v]$ due SVL normati. Si indica con*

$$B(T, V) := \{A : T \to V \mid \ A \text{ operatore lineare limitato}\} \tag{3.5}$$

lo SVL normato formato da tutti gli operatori lineari limitati da T a V. In particolare, si definisce

$$B(H) := \{T : H \to H \mid T \text{ è lineare}\} \in \mathrm{End}(H). \tag{3.6}$$

insieme degli endomorfismi lineari di H.

La norma di un operatore può adesso essere introdotta nella seguente maniera.

Definizione 3.9. *Sia ora $T : V \to W$ un operatore lineare. Si definisce una norma per l'operatore:*

$$\|T\|_{B(V,W)} = \sup_{x \in V \setminus \{0\}} \frac{\|T\,x\|_W}{\|x\|_V} = \sup_{\|x\|_V = 1} \|T\,x\|_W \tag{3.7}$$

che è proprio la norma che rende $B(T, V)$ normato.

Il fatto che l'Equazione 3.7 sia effettivamente una norma discende dal fatto che le funzioni $\|\cdot\|_V$ e $\|\cdot\|_W$ siano esse stesse delle norme nei rispettivi spazi vettoriali.

Se $\|T\| = +\infty$, l'operatore viene detto **illimitato**, altrimenti l'operatore è chiaramente limitato nel senso della definizione precedente. Infatti, non è difficile mostrare che

Proposizione 3.2. *Se $\|T\|_{B(V,W)} < +\infty$, allora $\|T\,x\|_W \leq \|T\|_{B(V,W)}\,\|x\|_V$, ovvero l'operatore è limitato anche nel senso della definizione precedentemente data.*

A questo punto, con l'ausilio di (3.7), si può mostrare che

Proposizione 3.3. *La composizione di operatori continui è anch'essa un operatore continuo.*

Dimostrazione. Dati due operatori $S : V \to W$, $T : W \to Z$ continui, si deve mostrare che $TS : V \to$

Z è limitato, ovvero continuo.

$$\begin{aligned}\|TS\|_{B(V,Z)} &= \sup_{\|x\|_V=1} \|T(S(x))\|_Z \le \sup_{\|x\|_V=1} \|T\|_{B(W,Z)} \|S(x)\|_W \\ &\le \sup_{\|x\|_V=1} \|T\|_{B(W,Z)} \|S\|_{B(V,W)} \|x\|_V = \|T\|_{B(W,Z)} \|S\|_{B(V,W)} = M \\ &\therefore \|TS\|_{B(V,Z)} \le M \in \mathbb{R}\end{aligned} \tag{3.8}$$

che dimostra la proprietà richiesta. □

Esempio 3.6. Alcuni esempi di operatori illimitati sono:

1. nello spazio $V = \mathscr{L}^2(-\pi, \pi)$, presa $\left\{ \varphi_k(x) = \frac{e^{ikx}}{\sqrt{2\pi}} \right\}$ come base ortonormale, si consideri l'operatore $T = \frac{d}{dx}$. Si ha

$$\|T\| = \sup_{\|f\|_V=1} \left\| \frac{df}{dx} \right\|_V \ge \sup_{k\in\mathbb{Z}} \left\| \frac{d\varphi}{dx} \right\|_V = \sup_{k\in\mathbb{Z}} \|ik\varphi_k\| = \sup_{k\in\mathbb{Z}} |k| = +\infty \tag{3.9}$$

   Allora, l'operatore T è illimitato;

2. nella trattazione dell'oscillatore armonico quantistico, si perviene a dover studiare gli autovalori e gli autostati dell'operatore hamiltoniano $\hat{\mathcal{H}}$, definito in questo caso come

$$\hat{\mathcal{H}} = \frac{1}{2}\left(-\frac{d^2}{dx^2} + x^2 \right) \tag{3.10}$$

   che ha autovalori, ovvero le energie permesse, pari a $E_n = n + 1/2$ (in opportune unità di misura) e autostati unitari che vengono indicati con $\{\varphi_n\}$. L'operatore $\hat{\mathcal{H}}$ risulta però non limitato:

$$\|\hat{\mathcal{H}}\| = \sup_{\|f\|=1} \|\hat{\mathcal{H}} f\| \ge \sup_{n\in\mathbb{N}} \|\hat{\mathcal{H}}\varphi_n\| = \sup_n \|E_n \varphi_n\| = \sup_n E_n = +\infty \tag{3.11}$$

   Questo avviene perché l'insieme delle energie permesse non è limitato superiormente. Non è difficile mostrare che se tale insieme è limitato, allora anche l'operatore è limitato.

Se A è un operatore illimitato, allora si dimostra che $\mathrm{dom}(A) \subset H$, ovvero non può coincidere con tutto lo spazio di partenza. Questo può creare problemi quando ad esempio si lavora con potenze di operatori illimitati, in quanto essere potrebbero risultare mal definite. Per operatori limitati, invece, che possono essere definiti su tutto lo spazio di partenza questo problema non si pone.
In spazi di Hilbert, si può esprimere la (3.7) in termini del prodotto scalare.

Proposizione 3.4. *Se $T \in B(H)$, allora si dimostra che*

$$\|T\|_{B(V,W)} = \sup_{\|x\|_V=1} \|Tx\|_W = \sup_{\|f\|=\|g\|=1} |\langle Tf, g\rangle| = \sup_{\|f\|=\|g\|=1} |\langle f, Tg\rangle|. \tag{3.12}$$

Dimostrazione. Usiamo il metodo *del sandwich*.

$$\begin{aligned} \sup |\langle Tf, g\rangle| &\leq \sup \|Tf\|\|g\| = \sup \|Tf\| \\ \sup |\langle Tf, g\rangle| &\geq \sup |\langle Tf, Tf\rangle| \frac{1}{\|Tf\|} = \sup \|Tf\| \end{aligned} \tag{3.13}$$

□

3.2.3 Inverso e aggiunto di un operatore

Dato $T : V \to W$, in alcuni casi è possibile definire l'inverso di T.

Definizione 3.10. *L'*inverso *di T è l'operatore $T^{-1} : W \to V$ tale che $TT^{-1} = T^{-1}T = \mathbb{1}$.*

Non è difficile mostrare che

Proposizione 3.5. *Se T ammette inverso, allora esso è unico.*

Dimostrazione. Si supponga che esistano due operatori X, Y tali che $TX = XT = \mathbb{1} \wedge TY = YT = \mathbb{1}$, allora $X = X\mathbb{1} = X(TY) = (XT)Y = \mathbb{1}Y = Y$. Questo ne prova l'unicità. □

Si può introdurre anche un altro importantissimo operatore.

Definizione 3.11. *Dato $A : H \to H$, si definisce $A^\dagger$ detto* **aggiunto** *di A come l'operatore tale che*

$$\left\langle A^\dagger f, g\right\rangle = \langle f, Ag\rangle \quad \forall f, g \in H \tag{3.14}$$

A differenza dell'inverso, per un operatore è sempre definito l'aggiunto. Bisogna notare però che la definizione (3.14) usa esplicitamente il fatto che l'operatore agisca su uno spazio di Hilbert. Se così non fosse, si dovrebbe trovare una definizione alternativa che non faccia uso in alcun modo del prodotto scalare. Alcune proprietà dell'aggiunto sono facilmente dimostrabili.

Proposizione 3.6. *Dato un operatore $A : H \to H$, si ha*

1. $A^{\dagger\dagger} = A$;
2. $(\lambda A + \mu B)^\dagger = \bar{\lambda} A^\dagger + \bar{\mu} B^\dagger$;
3. $(AB)^\dagger = B^\dagger A^\dagger$.

Dimostrazione.

1. $\left\langle (A^\dagger)^\dagger f, g\right\rangle = \left\langle f, A^\dagger g\right\rangle = \overline{\left\langle A^\dagger g, f\right\rangle} = \overline{\langle g, Af\rangle} = \langle Af, g\rangle \Rightarrow A^{\dagger\dagger} = A$
2. discende dalla linearità del prodotto scalare;
3. $\left\langle (AB)^\dagger f, g\right\rangle = \langle f, (AB)g\rangle = \langle f, A(Bg)\rangle = \left\langle A^\dagger f, Bg\right\rangle = \left\langle B^\dagger(A^\dagger f), g\right\rangle = \left\langle (B^\dagger A^\dagger) f, g\right\rangle \Rightarrow (AB)^\dagger = B^\dagger A^\dagger$.

□

Esempio 3.7. Si analizzino alcuni esempi di operatori aggiunti, banali e non:

1. *operatore nullo*: $\left\langle O^\dagger f, g\right\rangle = \langle f, Og\rangle = 0 \Rightarrow O^\dagger = O$;
2. *operatore identità*: $\left\langle \mathbb{1}^\dagger f, g\right\rangle = \langle f, \mathbb{1}g\rangle = \langle f, g\rangle \Rightarrow \mathbb{1}^\dagger = \mathbb{1}$;

3. dati $e_1, e_2 \in H$ fissati, $Pf = \langle e_1, f\rangle e_2$. Allora l'aggiunto $P^\dagger$ si ottiene da $\langle P^\dagger f, g\rangle = \langle f, Pg\rangle = \langle f, \langle e_1, g\rangle e_2\rangle = \langle e_1, g\rangle \langle f, e_2\rangle = \langle\langle e_2, f\rangle e_1, g\rangle$. Quindi $P^\dagger f = \langle e_2, f\rangle e_1$;

4. $A = \frac{d}{dx}$. Allora $\langle A^\dagger f, g\rangle = \left\langle f, \frac{dg}{dx}\right\rangle = \int_{\mathbb{R}} \overline{f(x)}\left(\frac{dg}{dx}\right) dx = \overline{f(x)}g(x)|_{\mathbb{R}} - \left\langle \frac{d}{dx}f, g\right\rangle = \left\langle -\frac{d}{dx}f, g\right\rangle \implies A^\dagger = -A \Leftrightarrow (\frac{d}{dx})^\dagger = -\frac{d}{dx}$[a].

5. $Af = a(x)f(x)$, con $a(x)$ funzione fissata, tale che $\int_{\mathbb{R}} |a(x)\, f(x)|^2\, dx < +\infty$. Una condizione sufficiente è chiedere che $a(x)$ sia essenzialmente limitata. L'operatore A è detto **operatore di moltiplicazione**. Si ha

$$\left\langle A^\dagger f, g\right\rangle = \langle f, Ag\rangle = \langle f, ag\rangle = \int_{\mathbb{R}} \overline{f(x)}[a(x)g(x)]\, dx$$
$$= \int_{\mathbb{R}} \overline{\overline{a(x)}f(x)}g(x)\, dx = \left\langle \overline{a(x)}f(x), g(x)\right\rangle$$

Allora $A^\dagger f = \overline{a(x)}f$.

[a] La cancellazione è possibile grazie a un teorema che afferma che se una funzione $g(x)$ e la sua derivata $g'(x)$ appartengono a $\mathscr{L}^2(\mathbb{R})$, allora $\lim_{x,\pm\infty} g(x) = 0$.

E se si volesse esprimere la condizione (3.14) per operatori che mappano uno spazio di Hilbert in un altro spazio di Hilbert? Questa situazione è di interesse pratico in molti casi. Ad esempio, dato uno spazio di Hilbert H_1 infinito dimensionale e un insieme totale di vettori ortonormali $\mathcal{F}_\varphi = \{\varphi_i\}$, l'operatore

$$T_\varphi f = \{\langle \varphi_i, f\rangle\, \forall\, i \in \mathbb{N}\} \tag{3.15}$$

è un operatore che restituisce un vettore in $\ell^2(\mathbb{N})$. Infatti, grazie all'uguaglianza di Parseval $\sum_n |\langle \varphi_i, f\rangle|^2 = \|f\|^2 < +\infty$. Quindi $T_\varphi : H_1 \to \ell^2(\mathbb{N})$. In una situazione simile, la condizione (3.14) andrebbe scritta considerando che nei due spazi in generale è definito un diverso prodotto scalare.

Definizione 3.12. *Per un operatore $A : H_1 \to H_2$, si chiama* **aggiunto** *di A l'operatore $A^\dagger$ tale che*

$$\langle A^\dagger f, g\rangle_1 = \langle f, Ag\rangle_2 \tag{3.16}$$

L'aggiunto di un operatore si comporta bene rispetto alla norma definita in (3.7). Valgono infatti le seguenti proprietà:

Proposizione 3.7. *Dato un operatore $A : H \to H$, si ha che*

1. $\|A^\dagger\|_{B(H)} = \|A\|_{B(H)}$;
2. $\|A^\dagger A\|_{B(H)} = \|A\|^2_{B(H)}$, *detta proprietà C^* (C star).*

Dimostrazione. 1. $\|A\|_{B(H)} = \sup |\langle A f, g\rangle|$
$\Rightarrow \|A^\dagger\|_{B(H)} = \sup_{\|f\|=\|g\|=1} |\langle A^\dagger f, g\rangle| = \sup_{\|f\|=\|g\|=1} |\langle f, A g\rangle| = \|A\|_{B(H)}$.

2. $\|A^\dagger A\|_{B(H)} \le \|A^\dagger\|_{B(H)}\|A\|_{B(H)} = \|A\|^2_{B(H)}$ per le proprietà precedentemente dimostrate.

Inoltre, $\|A\|^2_{B(H)} = \left(\sup_{\|f\|=1} \|A f\|\right)^2 = \sup_{\|f\|=1} \|A f\|^2 = \sup_{\|f\|=1} |\langle A f, A f\rangle|$

$= \sup_{\|f\|=1} \left|\langle A^\dagger A f, f\rangle\right| \leq \sup \|A^\dagger A f\| \|f\| = \|A^\dagger A\|_{B(H)}$, e quindi vale l'uguaglianza. □

3.3 Classi notevoli di operatori

Si introducono adesso alcune classi notevoli di operatori, utilissimi nelle applicazioni.

Definizione 3.13. *H è detto* **autoaggiunto** *("o hermitiano"[a]) se $H = H^\dagger$. U è detto* **unitario** *se $UU^\dagger = U^\dagger U = \mathbb{1}$ ovvero se $U^{-1} = U^\dagger$. P è detto un* **proiettore** *se $P^2 = P$. Il proiettore risulterà* **ortogonale** *se $P^2 = P = P^\dagger$.*

[a] In fisica sono usati come sinonimi, ma sono due concetti distinti in matematica. Nel seguito diremo sempre *autoaggiunto.*

Un operatore di proiezione, se autoaggiunto, è detto ortogonale perché $f = \mathbb{1} f = (\mathbb{1} - P + P) f = (\mathbb{1} - P) f + P f$, da cui $\langle (\mathbb{1} - P) f, P f\rangle = \langle P^\dagger (\mathbb{1} - P) f, f\rangle = \langle P (\mathbb{1} - P) f, f\rangle = 0$ perché $P^2 = P$. Quindi, un operatore siffatto scompone il vettore in due insiemi di componenti facenti parte di sottospazi ortogonali. Analizziamo in particolare il caso degli operatori unitari e autoaggiunti.

3.3.1 Operatori unitari

È noto che U è unitario se $U^{-1} = U^\dagger$. Operatori del genere sono detti unitari perché la loro azione non altera il prodotto scalare, e quindi la norma

$$\langle Uf, Ug\rangle = \left\langle U^\dagger U f, g\right\rangle = \langle f, g\rangle \tag{3.17}$$

Una proprietà importante è la seguente.

Proposizione 3.8. *Siano $\mathcal{F}_\varphi = \{\varphi_i\}$ una base totale di H e U un operatore unitario. $\mathcal{F}_\psi = \{\psi_i = U\varphi_i\}$ è una base totale dello spazio di Hilbert H.*

Dimostrazione. Data una qualsiasi f tale che $\langle f, \psi_i\rangle = 0$, allora

$$0 = \langle f, \psi_i\rangle = \langle f, U\varphi_i\rangle = \left\langle U^\dagger f, \varphi_i\right\rangle \Leftrightarrow U^\dagger f = 0 \Leftrightarrow U^{-1} f = 0. \tag{3.18}$$

Poiché U è invertibile, allora $f = 0$ e quindi la base è totale. □

Un esempio di operatore unitario è l'operatore di **traslazione**, definito come $T_a f = f(x-a)$, con $f \in \mathscr{L}^2(\mathbb{R})$ e $a > 0$. È ovvio che $T_a^{-1} f = f(x+a)$. Invece l'aggiunto è

$$\left\langle T_a^\dagger f, g\right\rangle = \langle f, T_a g\rangle = \int_{\mathbb{R}} \overline{f(x)} T_a g \, \mathrm{d}x = \int_{\mathbb{R}} \overline{f(x)} g(x-a) \, \mathrm{d}x = \int_{\mathbb{R}} \overline{f(y+a)} g(y) \, \mathrm{d}y$$

con $y := x - a$. Da ciò, $T_a^\dagger f = f(x+a)$ e quindi l'operatore è unitario.

Esempio 3.8. Affinché un operatore possa essere unitario, esso deve intanto essere invertibile. Capita così che operatori che conservano la norma, falliscono ad essere unitari a causa della loro non-invertibilità. Un esempio di tale situazione è l'operatore di **right-shift**

$$Ra = R(a_1, a_2, a_3, \dots) = (0, a_1, a_2, \dots) \tag{3.19}$$

È ovvio che $\|Ra\| = \|a\|$, ma R^{-1} non esiste. Infatti, sia L l'operatore di **left-shift** definito come

$$La = L(a_1, a_2, a_3, \dots) = (a_2, a_3, \dots), \tag{3.20}$$

candidato naturale ad essere l'inverso di R. È evidente che in questo caso $RL\,a \neq a$ in quanto

$$RLa = R(a_2, a_3, \dots) = (0, a_2, a_3, \dots) \neq a, \tag{3.21}$$

indi per cui $L \neq R^{-1}$.

3.3.2 Operatori autoaggiunti

Per gli operatori autoaggiunti vale la seguente importante caratterizzazione:

Teorema 3.1. *Se H è autoaggiunto e λ è un suo autovalore, allora $\lambda \in \mathbb{R}$. Inoltre, autovettori relativi ad autovalori distinti sono ortogonali tra loro.*

Dimostrazione. Se λ è un autovalore di $H \Rightarrow \langle f, Hf\rangle = \lambda \langle f, f\rangle \Rightarrow \lambda = {}^{\langle f,Hf\rangle}/_{\langle f,f\rangle}$. Ma $\overline{\langle f, Hf\rangle} = \overline{\langle H^\dagger f, f\rangle} = \langle f, H^\dagger f\rangle = \langle f, Hf\rangle \Rightarrow \langle f, Hf\rangle, \lambda \in \mathbb{R}$. Questo prova la prima parte.
Consideriamo due autovettori f, g relativi ad autovalori diversi λ, μ. Allora, $\langle g, Hf\rangle = \lambda \langle g, f\rangle$ e $\overline{\langle f, Hg\rangle} = \mu \overline{\langle f, g\rangle} \Rightarrow \langle Hg, f\rangle = \langle g, Hf\rangle = \mu \langle g, f\rangle$ da cui $(\lambda - \mu)\langle g, f\rangle = 0 \Leftrightarrow \langle g, f\rangle = 0$. □

Se l'operatore H non è autoaggiunto, viene meno *la sicurezza* che i suoi autovalori siano reali. In questo caso, condizione sufficiente per poter affermare che autovettori relativi ad autovalori differenti sono ortogonali tra loro è chiedere che $\lambda \neq \bar{\mu}$.

3.4 Regole di commutazione e anticommutazione canonica

Introduciamo adesso brevemente due classi di operatori, detti rispettivamente *fermionici* e *bosonici*, che risultano definiti a partire da una opportuna regola di (anti)commutazione. Il loro utilizzo è cruciale nella descrizione di innumerevoli sistemi in meccanica quantistica.

3.4.1 CAR: Canonical Anticommutation Relations

Le **regole di anticommutazione canonica** si applicano ad un operatore $b : H_f \to H_f$, dove il pedice f sta per *fermionico*. Le regole sono le seguenti:

$$\{b, b^\dagger\} = b\,b^\dagger + b^\dagger\,b = \mathbb{1} \tag{3.22}$$
$$\{b, b\} = bb + bb = 0 \tag{3.23}$$

che implicano $b^2 = (b^\dagger)^2 = 0$. Si può ricavare che tale operatore agisce su spazi di Hilbert a dimensione 2. Per argomentare questo, partiamo dalla seguente considerazione. Per tale operatore esiste sempre un vettore φ_0 tale che $b\varphi_0 = 0$. Se così non fosse, sarebbe possibile costruire il vettore $g = b\,f$ tale che $b\,g \not{0}$. Ma al contempo risulta che $b\,g = b^2 f = 0$, in chiara contraddizione con quanto detto prima.

Indichiamo adesso con $\varphi_1 = b^\dagger\,\varphi_0$. Allora $\langle \varphi_1, \varphi_0\rangle = \langle b^\dagger\,\varphi_0, \varphi_0\rangle = \langle \varphi_0, b\,\varphi_0\rangle = 0$, ovvero gli stati sono mutualmente ortogonali. Poiché $(b^\dagger)^n\,\varphi_0 = 0 \;\; n \geq 1$, allora lo spazio ha dimensione 2 e $\mathcal{F}_\varphi = \{\varphi_0, \varphi_1\}$ risulta esserne una base totale.

L'azione di b su φ_1 restituisce

$$b\,\varphi_1 = b\,b^\dagger\,\varphi_0 = (\mathbb{1} - b^\dagger\,b)\varphi_0 = \varphi_0 - b^\dagger\,\cancel{(b\,\varphi_0)} = \varphi_0. \tag{3.24}$$

Nel contesto di un'interpretazione particellare, si può interpretare questo risultato dicendo che b agisce come un *operatore di annichilazione* che distrugge una particella restituendo lo *stato di vuoto* φ_0 del sistema. Al contrario, $b^\dagger$ è detto essere un *operatore di creazione* che crea una particella (o eccitazione) portando il sistema nello stato φ_1.

Il conteggio delle *particelle* nei diversi stati è possibile tramite l'introduzione dell'operatore $N := b^\dagger b$, detto *operatore numero*. Si ha infatti che

$$N\varphi_0 = 0 \tag{3.25}$$

$$N\varphi_1 = b^\dagger\, b\, \varphi_1 = b^\dagger \varphi_0 = \varphi_1 \tag{3.26}$$

relazioni che si possono anche interpretare dicendo che φ_0 e φ_1 sono autostati di N, di autovalori rispettivamente 0 e 1. Non è difficile verificare che N è un operatore autoaggiunto, e questo è concorde col fatto che i suoi autovalori siano reali. Inoltre, $N^2 = N$, quindi N è un proiettore ortogonale.

Una possibile rappresentazione matriciale dell'operatore b e del suo aggiunto è la seguente

$$b = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad b^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{3.27}$$

Non è difficile verificare che questa scelta di b verifica le regole di anticommutazione canonica. L'operatore numero assume la forma

$$N = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.28}$$

che è una matrice diagonale avente gli autovalori sulla diagonale principale. Un calcolo diretto mostra che il vettore φ_0 è nella forma

$$\varphi_0 = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \quad \alpha \in \mathbb{C} \tag{3.29}$$

Si può prendere per semplicità $\alpha = 1$ e si ottiene ovviamente

$$\varphi_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \varphi_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.30}$$

La rappresentazione data di b non è unica ed è dettata da un criterio di semplicità. Tecnicamente, qualsiasi matrice unitariamente equivalente a quelle date fornirebbe la corretta rappresentazione dell'operatore b e del suo aggiunto. Nello spazio delle matrici 2×2, le trasformazioni unitarie sono descrivibili tramite rotazioni di un angolo θ. Così utilizzando la nota espressione di una matrice di rotazione $U(\theta)$ sul piano, si ottiene

$$c = U(\theta)\, b\, U^{-1}(\theta) = \begin{pmatrix} \sin\theta\cos\theta & \cos^2\theta \\ -\sin^2\theta & -\sin\theta\cos\theta \end{pmatrix} \tag{3.31}$$

che soddisfa ancora le regole di anticommutazione canoniche.

3.4.2 CCR: Canonical Commutation Relations

Le **regole di commutazione canonica** sono soddisfatte da operatori definiti in generale su spazi di Hilbert infinito-dimensionali, detti spazi di Hilbert *bosonici*. La regola che definisce tali operatori è

$$\left[a, a^\dagger\right] = a\, a^\dagger - a^\dagger\, a = \mathbb{1}. \tag{3.32}$$

Anche qui, si può dimostrare l'esistenza di uno **stato ground** $\varphi_0 \neq 0$, ovvero tale che $a\varphi_0 = 0$. Visto che non è posta alcuna condizione sulle potenze di a, è possibile definire in generale

$$\varphi_n := \frac{(a^\dagger)^n}{\sqrt{n!}}\varphi_0. \tag{3.33}$$

Questa definizione produce sempre stati diversi dal vettore nullo. Infatti, si può dimostrare che $\langle \varphi_n, \varphi_m \rangle = \delta_{n,m}$. Per mostrare questo, si definisca l'*operatore numero* $N = a^\dagger a$. Si può far vedere per induzione che $N\varphi_n = n\varphi_n$. Infatti $N\varphi_0 = 0$ ovviamente. Supposta vera la proprietà per n, si ha

$$\begin{aligned} N\varphi_{n+1} &= a^\dagger a \left[\frac{(a^\dagger)^{n+1}}{\sqrt{(n+1)!}}\varphi_0\right] = a^\dagger a \left[\frac{a^\dagger}{\sqrt{n+1}}\frac{(a^\dagger)^n}{\sqrt{n!}}\varphi_0\right] \\ &= \varphi_{n+1} + \frac{a^\dagger}{\sqrt{n+1}}\underbrace{a^\dagger a}_{N}\varphi_n = (n+1)\varphi_{n+1} \end{aligned} \tag{3.34}$$

Quindi i vari φ_n sono autostati dell'operatore autoaggiunto N. Allora, sono ortogonali tra loro. Si mostri ora che $\|\varphi_n\| = 1$. φ_0 può essere normalizzato in maniera da avere norma unitaria. Supposta vera la proprietà per n, mostriamola vera per $n+1$:

$$\begin{aligned} \|\varphi_{n+1}\|^2 = \langle \varphi_{n+1}, \varphi_{n+1} \rangle &= \left\langle \frac{a^\dagger}{\sqrt{n+1}}\varphi_n, \frac{a^\dagger}{\sqrt{n+1}}\varphi_n \right\rangle = \frac{1}{n+1}\left\langle a^\dagger\varphi_n, a^\dagger\varphi_n \right\rangle = \\ &= \frac{1}{n+1}\left\langle \varphi_n, a\,a^\dagger\varphi_n \right\rangle = \frac{1}{n+1}\langle \varphi_n, \varphi_n \rangle + \frac{1}{n+1}\langle \varphi_n, N\varphi_n \rangle = 1 \end{aligned} \tag{3.35}$$

Questo completa la dimostrazione che $\langle \varphi_n, \varphi_m \rangle = \delta_{n,m}$.

Tra le scritture precedenti, è anche contenuto il fatto che $a^\dagger\varphi_n = \sqrt{n+1}\varphi_{n+1}$. Quindi anche qui $a^\dagger$ ha il ruolo, a meno di un fattore di normalizzazione, di un operatore di creazione. Invece, applicando a alla precedente relazione si ottiene $a\,\varphi_{n+1} = \sqrt{n+1}\varphi_n$; allora a è un operatore di annichilazione.

Nel caso fermionico, gli operatori sono sempre limitati in quanto lo spazio è a dimensione finita. Nel caso bosonico invece non è sempre detto. Infatti ad esempio

$$\|N\|_{B(H)} = \sup_{\|f\|=1} \|Nf\| \geq \sup_n \|N\varphi_n\| = \sup_n n\|\varphi_n\| = +\infty \tag{3.36}$$

e quindi l'operatore numero nel caso bosonico è illimitato.

Una rappresentazione dell'operatore a è

$$a = \frac{1}{\sqrt{2}}\left(x + \frac{\mathrm{d}}{\mathrm{d}x}\right), \; a^\dagger = \frac{1}{\sqrt{2}}\left(x - \frac{\mathrm{d}}{\mathrm{d}x}\right) \tag{3.37}$$

Non è difficile verificare che $[a, a^\dagger]f(x) = f(x)$. A partire da questa rappresentazione, si trova che $\varphi_0 = c_0 \exp\{-x^2/2\}$. Un'alternativa esotica per rappresentare a è mediante la matrice infinita

$$a = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{2} & 0 & \dots \\ 0 & 0 & 0 & \sqrt{3} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.38}$$

3.5 Principio di indeterminazione di Heisenberg

Per operatori autoaggiunti a commutatore non nullo vale un'importante disuguaglianza. Il punto di partenza sono due operatori autoaggiunti A, B tali che $[A, B] = i\,C$, dove C è un operatore. Su C non viene fatta alcuna richiesta di regolarità, ma si ricava facilmente che esso è autoaggiunto. Infatti

$$C = -i(AB - BA) = -i[A, B] = i(BA - AB) = C^{\dagger}. \tag{3.39}$$

D'ora in poi, si lavorerà con stati $\varphi \in H$ tale che $\|\varphi\| = 1$, ovvero stati normalizzati.

Definizione 3.14. *Si definisce il* **valore medio** *di X nello stato φ la grandezza $\langle X\rangle_\varphi = \langle \varphi, X\varphi\rangle$.*

Definizione 3.15. *Si definisce la* **varianza** *di X nello stato φ come $\Delta^2 X = \left\langle X^2\right\rangle_\varphi - \langle X\rangle_\varphi^2 = \left\langle (X - \langle X\rangle_\varphi)^2\right\rangle_\varphi$. La* **deviazione standard** *nello stato φ è definita come $\Delta X = \sqrt{\Delta^2 X}$.*

Si supponga che X sia un operatore autoaggiunto. Allora sicuramente:

1. $\langle X\rangle \in \mathbb{R}$. Infatti $\overline{\langle \varphi, X\varphi\rangle} = \langle X\varphi, \varphi\rangle = \langle \varphi, X\varphi\rangle$;
2. $\langle (X - \langle X\rangle)^2\rangle = \langle \varphi, (X - \langle X\rangle)^2\varphi\rangle = \langle (X - \langle X\rangle)\varphi, (X - \langle X\rangle)\varphi\rangle = \|(X - \langle X\rangle)\varphi\|^2$.

Allora, si possono definire due nuovi operatori, $\hat{A} = A - \langle A\rangle$ e $\hat{B} = B - \langle B\rangle$, evidentemente autoaggiunti e tali che $[\hat{A}, \hat{B}] = [A, B]$. Allora, detto $X_\alpha = \hat{A} + i\alpha\hat{B}$, si ha che $0 \le \|X_\alpha\varphi\|^2 = \langle (\hat{A} + i\alpha\hat{B})\,\varphi, (\hat{A} + i\alpha\hat{B})\,\varphi\rangle = \|\hat{A}\,\varphi\|^2 + |\alpha|^2\,\|\hat{B}\,\varphi\|^2 + i\,\alpha\,\langle \varphi, [\hat{A}, \hat{B}]\,\varphi\rangle = |\alpha|^2\,\|\hat{B}\,\varphi\|^2 - \alpha\,\langle C\rangle + \|\hat{A}\,\varphi\|^2$. Affinché questa disuguaglianza valga $\forall\,\alpha \in \mathbb{R}$, il discriminante dell'equazione di secondo grado associata deve essere non positivo. Questo comporta che $\|\hat{A}\varphi\|\,\|\hat{B}\varphi\| \ge \frac{1}{2}|\langle C\rangle|$. Poiché però evidentemente $\|\hat{A}\varphi\| = \Delta A$, così anche per B, si ha infine che

$$\frac{1}{2}\,|\langle C\rangle| \le \Delta A\,\Delta B \tag{3.40}$$

Questa disuguaglianza, valida per qualsiasi coppia di operatori autoaggiunti, ha una particolare applicazione in meccanica quantistica. Infatti, poiché gli operatori autoaggiunti $\hat{x}$ e $\hat{p}$ sono tali che $[\hat{x}, \hat{p}] = \hbar\mathbb{1}$, allora

$$\Delta\hat{x}\,\Delta\hat{p} \ge \frac{\hbar}{2} \tag{3.41}$$

La relazione di indeterminazione posizione-impulso è nota come **principio di indeterminazione di Heisenberg**. Si ottiene la minima indeterminazione ad esempio per *stati coerenti* .

Capitolo 4

Trasformate integrali: Fourier, Laplace

In questo capitolo analizziamo due importanti classi di trasformate integrali, ovvero le trasformate di Fourier e di Laplace. Iniziamo il capitolo introducendo la definizione e le proprietà del prodotto di convoluzione.

4.1 Introduzione: il prodotto di convoluzione

Diamo qui la definizione di una maniera di *combinare* due funzioni f e g qualsiasi tramite un opportuno integrale.

Definizione 4.1. *Si definisce* **convoluzione** *tra due funzioni f e g tramite il seguente integrale:*

$$(f \star g)(x) := \int_{\mathbb{R}} f(y)g(x-y)\,\mathrm{d}y \tag{4.1}$$

Alcune proprietà sono riassunte nella seguente

Proposizione 4.1. *La convoluzione tra due funzioni f, g è un'operazione commutativa: $(f \star g)(x) = (g \star f)(x)$. Inoltre, supposto che f non sia derivabile e g sia derivabile, allora $\frac{\mathrm{d}}{\mathrm{d}x}(f \star g)(x) = \int_{\mathbb{R}} f(x)g'(x-y)\,\mathrm{d}y$.*

L'ultima proprietà in particolare permette in qualche modo di *derivare* anche le funzioni non derivabili.

4.2 Trasformata di Fourier

Uno sviluppo in serie di Fourier di una funzione $f(x)$ è possibile solo se essa è, in generale, $T-$*periodica*. In quel caso, la serie fornirà una rappresentazione del prolungamento periodico di f e lo spazio ambiente su cui considerare f sarà $\mathscr{L}^2(-\frac{\pi}{T}, \frac{\pi}{T})$. Per *sviluppare* funzioni non periodiche, sarà necessario passare da una rappresentazione in serie ad una **rappresentazione integrale**.

Definizione 4.2. *Si definisce* **trasformata di Fourier** *di una funzione f l'integrale*

$$\mathcal{F}[f](p) = \hat{f}(p) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)\mathrm{e}^{-ipx}\,\mathrm{d}x \quad \forall p \in \mathbb{R} \tag{4.2}$$

Tale trasformata integrale può essere pensata come un operatore integrale lineare $\mathcal{F}$. Si pone allora l'annosa questione di studiare dominio e immagine dell'operatore $\mathcal{F}$. Nelle seguenti sezioni verranno valutate alcune casistiche di interesse.

4.2.1 Operatore di Fourier per funzioni $\mathscr{L}^1$

Si dimostra facilmente che

Proposizione 4.2. *L'operatore $\mathcal{F}$ è definito da $\mathscr{L}^1(\mathbb{R})$ a $\mathscr{L}^\infty(\mathbb{R})$.*

Dimostrazione.

$$|\hat{f}(p)| = \frac{1}{\sqrt{2\pi}} \left| \int_{\mathbb{R}} f(x) \mathrm{e}^{-ipx}\, \mathrm{d}x \right| \leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x)| \underbrace{|\mathrm{e}^{-ipx}|}_{1} \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \|f\|_1.$$

Ciò implica, passando agli estremi superiori:

$$\left\|\hat{f}\right\|_\infty = \sup_{p\in\mathbb{R}} |\hat{f}(p)| \leq \frac{1}{\sqrt{2\pi}} \|f\|_1 < +\infty$$

$$\therefore \left\|\hat{f}\right\|_\infty < +\infty \implies \hat{f} \in \mathscr{L}^\infty$$

□

Indipendentemente dalla natura di f, la sua trasformata è una funzione continua.

Proposizione 4.3. *Se $\mathcal{F} : \mathscr{L}^1(\mathbb{R}) \to \mathscr{L}^\infty(\mathbb{R})$, allora $\mathcal{F}[f] = \hat{f}(p)$ è una funzione continua su tutto $\mathbb{R}$.*

Dimostrazione. $\hat{f}(p)$ è continua in $p_0 \iff \lim_{p,p_0} |\hat{f}(p) - \hat{f}(p_0)| = 0$. Ma verificare il limite è immediato:

$$\lim_{p,p_0} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x)||\mathrm{e}^{-ipx} - \mathrm{e}^{-ip_0x}|\, \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x)| \lim_{p,p_0} |\mathrm{e}^{-ipx} - \mathrm{e}^{-ip_0x}|\, \mathrm{d}x = 0$$

avendo usato il teorema della convergenza dominata per lo scambio di limite e integrale. Dall'arbitrarietà di p_0, segue la tesi. □

Definizione 4.3. *La* **trasformata di Fourier inversa** $\mathcal{F}^{-1}$ *è definita come*

$$\mathcal{F}^{-1}\left[\hat{f}\right](x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(p) \mathrm{e}^{ipx}\, \mathrm{d}p \quad \forall x \in \mathbb{R} \tag{4.3}$$

Tale trasformata spesso non è ben definita, in quanto se $\hat{f}(p) \in \mathscr{L}^\infty(\mathbb{R})$ non è detto che l'integrale in (4.3) converga per ogni x. Si dimostra però che richiedendo che $\hat{f} \in \mathscr{L}^1(\mathbb{R})$, si ha il seguente teorema, detto di **inversione locale**.

Teorema 4.1 (inversione locale). *Sia $f(x) \in \mathscr{L}^1(\mathbb{R})$ tale che $\hat{f} \in \mathscr{L}^1(\mathbb{R})$. Allora*

$$\mathcal{F}^{-1}\left[\hat{f}\right] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(p) \mathrm{e}^{ipx}\, \mathrm{d}x = f(x) \;\; q.o. \text{ in } \mathbb{R} \tag{4.4}$$

L'operatore $\mathcal{F}^{-1}$ è quindi l'inverso dell'operatore $\mathcal{F}$ soltanto per funzioni che stanno almeno nel sottospazio $\mathscr{L}^1 \cap \mathscr{L}^\infty$. Inoltre, non è difficile rendersi conto che f è anch'essa continua su tutto $\mathbb{R}$. L'inverso, insomma, esiste soltanto in situazioni molto particolari, ovvero se $f \in C^0(\mathbb{R}) \cap \mathscr{L}^1(\mathbb{R}) \cap \mathscr{L}^\infty(\mathbb{R})$.

Si introduce ora la seguente

Proposizione 4.4. *Se $f(x), g(x), \hat{f}(p) \in \mathscr{L}^1(\mathbb{R})$, allora $\exists \int_\mathbb{R} \overline{f(x)} g(x)\, \mathrm{d}x$ ed $\exists \int_\mathbb{R} \overline{\hat{f}(p)} \hat{g}(p)\, \mathrm{d}p$ e questi sono uguali.*

Dimostrazione. Per mostrare l'esistenza dei due integrali, si considera

$$\left| \int_\mathbb{R} \overline{\hat{f}(p)} \hat{g}(p)\, \mathrm{d}p \right| \leq \int_\mathbb{R} |\hat{f}(p)||\hat{g}(p)|\, \mathrm{d}p \leq \|g\|_\infty \int_\mathbb{R} |\hat{f}(p)|\, \mathrm{d}p \leq \|\hat{f}\|_1 \|\hat{g}\|_\infty < +\infty$$

viste le ipotesi. Analogamente,

$$\left| \int_\mathbb{R} \overline{f(x)} g(x)\, \mathrm{d}x \right| \leq \int_\mathbb{R} |f(x)||g(x)|\, \mathrm{d}x \leq \|f\|_\infty \int_\mathbb{R} |g(x)|\, \mathrm{d}x = \|f\|_\infty \|g\|_1 < +\infty$$

visto che per ipotesi $f \in \mathscr{L}^1 \cap \mathscr{L}^\infty$.
Questo mostra l'esistenza di ambedue gli integrali. L'uguaglianza si mostra calcolando

$$\begin{aligned}
\int_\mathbb{R} \overline{\hat{f}(p)} \hat{g}(p)\, \mathrm{d}p &= \int_\mathbb{R} \overline{\left(\frac{1}{\sqrt{2\pi}} \int_\mathbb{R} f(x) e^{-ipx}\, \mathrm{d}x \right)} \left(\frac{1}{\sqrt{2\pi}} \int_\mathbb{R} g(y) e^{-ipy}\, \mathrm{d}y \right) \mathrm{d}p \\
&= \int_\mathbb{R} \overline{f(x)} \int_\mathbb{R} g(y) \left(\frac{1}{2\pi} \int_\mathbb{R} e^{ip(x-y)}\, \mathrm{d}p \right) \mathrm{d}y\, \mathrm{d}x \\
&= \int_\mathbb{R} \overline{f(x)} \int_\mathbb{R} g(y) \delta(x-y)\, \mathrm{d}y\, \mathrm{d}x \\
&= \int_\mathbb{R} \overline{f(x)} g(x)\, \mathrm{d}x
\end{aligned}$$

sfruttando la rappresentazione integrale di Equazione 5.2 della δ di Dirac. □

Questo permette di introdurre un prodotto scalare in un sottospazio di $\mathscr{L}^1$.

4.2.2 Operatore di Fourier per funzioni in $\mathscr{L}^2$

Lo sviluppo della teoria richiede in questo caso un po' più di attenzione. In questo caso bisognerà introdurre alcune nozioni preliminari e dimostrare alcuni lemmi preparatori al fine di dimostrare un analogo della Proposizione 4.4.

Definizione 4.4. *Si chiama* spazio delle funzioni a decrescenza rapida *lo spazio*

$$\mathscr{S}(\mathbb{R}) = \left\{ f(x) \in C^\infty(\mathbb{R}) : \forall n, k \in \mathbb{N} \cup \{0\},\ \lim_{|x|, +\infty} |x|^n f^{(k)}(x) = 0 \right\} \tag{4.5}$$

In $\mathscr{S}(\mathbb{R})$ stanno quindi tutte le funzioni che vanno a 0, insieme a tutte le loro derivate, più velocemente di ogni potenza di x. Quindi, ad esempio, $f(x) = \frac{1}{1+x^2} \notin \mathscr{S}(\mathbb{R})$ mentre in generale $g_n(x) = p(x)e^{-x^{2n}} \in \mathscr{S}(\mathbb{R})$, dove $p(x)$ è un polinomio di grado qualsiasi. Data la definizione, è ovvio che $\mathscr{S}(\mathbb{R}) \subset \mathscr{L}^p(\mathbb{R})\ ,\quad 1 \leq p \leq +\infty$.
L'importanza di questo spazio, tra le altre cose, sta nella seguente:

Proposizione 4.5. *Ogni $f(x) \in \mathscr{L}^2(\mathbb{R})$ si può scrivere come limite di una successione di funzioni di $\mathscr{S}(\mathbb{R})$, ovvero $\mathscr{S}(\mathbb{R})$ è denso in $\mathscr{L}^2(\mathbb{R})$.*

Inoltre, si può mostrare che:

Proposizione 4.6. *Se $\varphi(x) \in \mathscr{S}(\mathbb{R})$, allora $\hat{\varphi}(p) \in \mathscr{S}(\mathbb{R})$.*

Dimostrazione. Poiché $\mathscr{S}(\mathbb{R}) \subset \mathscr{L}^1(\mathbb{R})$, allora sicuramente $\hat{\varphi}(p)$ esiste. Inoltre $\forall n \in \mathbb{N}$

$$\begin{aligned}\frac{d^n}{dp^n}[\hat{\varphi}(p)] &= \frac{1}{\sqrt{2\pi}}\frac{d^n}{dp^n}\int_{\mathbb{R}} \varphi(x)e^{-ipx}\,dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \varphi(x)\frac{\partial^n}{\partial p^n}e^{-ipx}\,dx = \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \underbrace{\varphi(x)(-ix)^n}_{=:\psi_n\in\mathscr{S}(\mathbb{R})} e^{-ipx}\,dx = \hat{\psi}_n(p)\end{aligned} \tag{4.6}$$

ovvero la derivata n-esima di $\hat{\varphi}(p)$ può essere vista come la trasformata di Fourier di una funzione $\psi_n(x)$ anch'essa in $\mathscr{S}(\mathbb{R})$, quindi è ben definita. Allora, sicuramente $\hat{\varphi}(p) \in C^\infty(\mathbb{R})$. Si mostri ora la tesi. Si ha

$$\begin{aligned}\lim_{|p|,+\infty} |p|^n\hat{\varphi}(p) &= \lim_{|p|,+\infty} \frac{1}{\sqrt{2\pi}}|p|^n\int_{\mathbb{R}} \varphi(x)e^{-ipx}\,dx \\ &= \lim_{|p|,+\infty} \frac{i^n}{\sqrt{2\pi}}\int_{\mathbb{R}} \varphi(x)\frac{\partial^n}{\partial x^n}\left[e^{-ipx}\right]dx\end{aligned} \tag{4.7}$$

Agendo per induzione su n, si può dimostrare che $|p^n\hat{\varphi}(p)| \leq \frac{1}{\sqrt{2\pi}}\underbrace{\int_{\mathbb{R}} |\varphi^{(n)}(x)|\,dx}_{M_n} = \frac{M_n}{\sqrt{2\pi}}$. Quindi $|\hat{\varphi}(p)| \leq \frac{M_n}{|p|^n\sqrt{2\pi}}$, ovvero va a zero più velocemente di ogni potenza di $|p|$. Poiché le derivate di $\hat{\varphi}(p)$ possono essere viste come trasformate di Fourier di funzioni in $\mathscr{S}(\mathbb{R})$, allora anche esse vanno a zero più velocemente di ogni potenza di $|p|$. Quindi $\hat{\varphi}(p) \in \mathscr{S}(\mathbb{R})$. □

Quindi, $\mathcal{F} : \mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R})$. Questo garantisce l'esistenza dall'antitrasformata per funzioni di $\mathscr{S}(\mathbb{R})$. Inoltre, poiché $\mathscr{S}(\mathbb{R}) \subset \mathscr{L}^1(\mathbb{R})$, allora vale la Proposizione 4.4 e quindi

$$\int_{\mathbb{R}} \overline{\varphi(x)}\psi(x)\,dx = \int_{\mathbb{R}} \overline{\hat{\varphi}(x)}\hat{\psi}(x)\,dx < +\infty \tag{4.8}$$

Inoltre, poiché $\mathscr{S}(\mathbb{R}) \subset \mathscr{L}^2(\mathbb{R})$, questi sono effettivamente prodotti scalari e si può scrivere che $\langle\varphi,\psi\rangle = \langle\hat{\varphi},\hat{\psi}\rangle\ \forall\varphi,\psi \in \mathscr{S}(\mathbb{R})$. Da questo segue immediatamente che $\|\varphi\| = \|\mathcal{F}[\varphi]\|\ ,\forall\varphi \in \mathscr{S}(\mathbb{R})$. Ma allora

$$\|\mathcal{F}\| = \sup_{\|\varphi\|=1} \|\mathcal{F}[\varphi]\| = \sup_{\|\varphi\|=1} \|\varphi\| = 1 \tag{4.9}$$

ovvero $\mathcal{F}$ è un operatore limitato, quindi continuo. Questo risulta molto utile perché implica che se $f_n \xrightarrow{\|\cdot\|} f$, allora $\mathcal{F}[f_n] \xrightarrow{\|\cdot\|} \mathcal{F}[f]$. Tale proprietà permette di estendere il dominio di $\mathcal{F}$ a tutto $\mathscr{L}^2(\mathbb{R})$, mostrando che

Teorema 4.2. *$\mathcal{F} : \mathscr{L}^2(\mathbb{R}) \to \mathscr{L}^2(\mathbb{R})$, ovvero $\mathcal{F}$ è un endomorfismo di $\mathscr{L}^2(\mathbb{R})$.*

Dimostrazione. Sia $f \in \mathscr{L}^2(\mathbb{R})$. Per la densità di $\mathscr{S}(\mathbb{R})$ in $\mathscr{L}^2(\mathbb{R})$ esiste una successione $\{f_n\} \subset \mathscr{S}(\mathbb{R})$ tale che $\|f - f_n\| \to 0$, ovvero $\forall \epsilon > 0\ \exists n_\epsilon > 0 : \|f_n - f_m\| < \epsilon\ \forall n, m > n_\epsilon$. Allora

$$\|\hat{f}_n - \hat{f}_m\| = \|\mathcal{F}[f_n] - \mathcal{F}[f_m]\| = \|\mathcal{F}[f_n - f_m]\| = \|f_n - f_m\| < \epsilon \tag{4.10}$$

visto che $f_n - f_m \in \mathscr{S}(\mathbb{R})$. Quindi anche $\{\hat{f}_n\}$ è di Cauchy e quindi convergente, ovvero $\exists \lim_{n,+\infty} \hat{f}_n(p) = \hat{f}(p)$. Dalla completezza di $\mathscr{L}^2$, segue che $\hat{f}(p) \in \mathscr{L}^2(\mathbb{R})$. □

Si osservi quindi che la trasformata per funzioni $\mathscr{L}^2$ è definita formalmente come limite della successione delle trasformate di funzioni $\mathscr{S}(\mathbb{R})$.

Proposizione 4.7. *Se $f, g \in \mathscr{L}^2(\mathbb{R})$, allora $\langle f, g\rangle = \left\langle \hat{f}, \hat{g}\right\rangle$.*

Dimostrazione. Siano $\{f_n\}_n$ e $\{g_n\}_n$ due successioni di funzioni tali che $\|f - f_n\| \to 0$ e $\|g - g_n\| \to 0$. Allora

$$\begin{aligned}\langle f, g\rangle &= \langle f - f_n + f_n, g - g_n + g_n\rangle \\ &= \langle f - f_n, g - g_n\rangle + \langle f - f_n, g_n\rangle + \langle f_n, g - g_n\rangle + \langle f_n, g_n\rangle\end{aligned} \tag{4.11}$$

Si analizzino i singoli contributi di (4.11):

$$\begin{aligned}|\langle f - f_n, g - g_n\rangle| &\le \|f - f_n\| \|g - g_n\| \to 0 \\ |\langle f - f_n, g_n\rangle| &\le \underbrace{\|f - f_n\|}_{\to 0} \underbrace{\|g_n\|}_{<+\infty} \to 0 \\ |\langle f_n, g - g_n\rangle| &\le \underbrace{\|f_n\|}_{<+\infty} \underbrace{\|g - g_n\|}_{\to 0} \to 0\end{aligned}$$

Allora $|\langle f, g\rangle - \langle f_n, g_n\rangle| \to 0$. Analogamente $\left|\left\langle \hat{f}, \hat{g}\right\rangle - \left\langle \hat{f}_n, \hat{g}_n\right\rangle\right| \to 0$. Ma poiché $\{f_n\}, \{g_n\} \subset \mathscr{S}(\mathbb{R})$, si ha $\langle f_n, g_n\rangle = \left\langle \hat{f}_n, \hat{g}_n\right\rangle$. Dall'unicità del limite, segue la tesi. □

Questo completa l'estensione della trasformata di Fourier a funzioni a quadrato sommabili. Questi risultati sono raccolti nel seguente teorema, dovuto a **Michel Plancherel**.[1]

Teorema 4.3 (di Plancherel). *La trasformata di Fourier di una funzione di $\mathscr{L}^2(\mathbb{R})$ sta in $\mathscr{L}^2(\mathbb{R})$. La sua antitrasformata restituisce quasi ovunque la funzione di partenza. Vale inoltre l'***uguaglianza di Parseval** *e l'*uguaglianza di Parseval generalizzata *(proprietà di preservazione della norma):*

$$\int_{\mathbb{R}} |f(x)|^2 \,\mathrm{d}x = \int_{\mathbb{R}} |\hat{f}(p)|^2 \,\mathrm{d}p \quad e \quad \int_{\mathbb{R}} \overline{f(x)} g(x) \,\mathrm{d}x = \int_{\mathbb{R}} \overline{\hat{f}(p)} \hat{g}(p) \,\mathrm{d}p \tag{4.12}$$

[1] M. Plancherel, *"Contribution à l'étude de la représentation dune fonction arbitraire par des intégrales définies"*, 1910, *Rendiconti del Circolo Matematico di Palermo*

Proprietà 4.1. *Alcune proprietà della trasformata di Fourier sono le seguenti:*

1. $\mathcal{F}\left[\frac{d^n}{dx^n}f\right] = (ip)^n \mathcal{F}[f]$;
2. $\mathcal{F}[x^n f] = i^n \frac{d^n}{dp^n}\mathcal{F}[f]$;
3. $\mathcal{F}[f_1 \star f_2] = \sqrt{2\pi}\hat{f}_1\hat{f}_2$;
4. $\mathcal{F}[f_1 f_2](p) = \frac{1}{\sqrt{2\pi}}(\mathcal{F}[f_1] \star \mathcal{F}[f_2])(p)$.

Dimostrazione. Le dimostrazioni delle proprietà (1) e (2) sono rintracciabili nello sviluppo dei calcoli precedentemente affrontato. Per la proprietà (3) invece, ricordando la definizione di convoluzione, si ha

$$\begin{aligned}\mathcal{F}[f_1 \star f_2] &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} f_1(\tau)f_2(x-\tau)d\tau\right)e^{-ipx}\,dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} f_1(\tau)f_2(y)d\tau\right)e^{-ip(\tau+y)}\,dy \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-ip\tau}f_1(\tau)d\tau\int_{\mathbb{R}} e^{-ipy}f_2(y)\,dy \\ &= \sqrt{2\pi}\hat{f}_1(p)\hat{f}_2(p)\end{aligned} \tag{4.13}$$

avendo posto $y := x - \tau$. La proprietà (4) non verrà dimostrata. □

In particolare, le ultime due proprietà affermano che l'operatore $\mathcal{F}$ di Fourier trasforma prodotti di convoluzione in prodotti ordinari, e viceversa.

Proposizione 4.8. *Valgono le seguenti proprietà:*

1. *se $f(x)$ è reale e pari, allora $\mathcal{F}[f]$ è reale e pari;*
2. *se $f(x)$ è reale e dispari, allora $\mathcal{F}[f]$ è puramente immaginaria e dispari.*

Infine, non è difficile mostrare che $\mathcal{F}$ è un operatore unitario. Infatti

$$\langle f, g\rangle = \left\langle \hat{f}, \hat{g}\right\rangle = \langle \mathcal{F}[f], \hat{g}\rangle = \left\langle f, \mathcal{F}^{\dagger}[\hat{g}]\right\rangle \iff g = \mathcal{F}^{\dagger}[\hat{g}] \iff \mathcal{F}^{\dagger} = \mathcal{F}^{-1} \tag{4.14}$$

Questo è in completo accordo con il fatto precedentemente ricavato che $\|\mathcal{F}\| = 1$. Riassumendo tutte le situazioni analizzate, si ha che

$$\mathcal{F} : \begin{cases} \mathscr{L}^1 \to \mathscr{L}^\infty \\ \mathscr{S} \to \mathscr{S} \\ \mathscr{L}^2 \to \mathscr{L}^2 \end{cases} \tag{4.15}$$

ovvero che, se una funzione $f(x) \in \mathscr{L}^1$, a seguito dell'applicazione dell'operatore $\mathcal{F}$ su di essa, la funzione $F[f] = \hat{f}(p) \in \mathscr{L}^\infty$, e così via. Bisogna dunque sempre controllare che i risultati ottenuti siano coerenti con tale schema.

Alcuni esempi

Esempio 4.1 (Norma di una funzione). Si calcoli la norma nel senso di $\mathscr{L}^2\left(\int_{\mathbb{R}} |\cdot|^2\right)$ della funzione complessa di variabile reale

$$f(x) = \frac{3}{(x - i\pi)^2}. \tag{4.16}$$

Il calcolo dell'integrale richiesto sarebbe di base non immediato. La richiesta può comunque essere soddisfatta, ricordando la Proposizione 4.7, ovvero che la trasformata di Fourier preserva la norma. Chiaramente, *si spera* di essere fortunati, ottenendo una funzione più semplice da integrare. Calcoliamo quindi

$$\mathcal{F}\left[f(x)\right](p) = 3\sqrt{2\pi} p e^{\pi p} u(-p) \tag{4.17}$$

ove $u(x) = 1$ se $x > 0$ e vale 0 altrove[a]. Concentriamoci adesso sul calcolo della norma

$$\|f(x)\|_2 = \left\|\hat{f}(p)\right\|_2 = \int_{-\infty}^{+\infty} \left|3\sqrt{2\pi} p e^{\pi p} u(-p)\right|^2 \mathrm{d}p = \int_{-\infty}^{0} \left|3\sqrt{2\pi} p e^{\pi p}\right|^2 \mathrm{d}p. \tag{4.18}$$

Quindi, dobbiamo calcolare $\int_{-\infty}^{0} p^2 e^{2\pi p} \mathrm{d}p$ che deve convergere ed essere positivo. Per risolverlo, possiamo utilizzare l'integrazione per parti due volte, ottenendo

$$\int_{-\infty}^{0} p^2 e^{2\pi p} \mathrm{d}p = \cancel{p^2 \frac{1}{2\pi} e^{2\pi p}\Big|_{-\infty}^{0}} - \int_{-\infty}^{0} \frac{1}{2\pi} e^{2\pi p} 2p \mathrm{d}p.$$

Il primo termine si annulla perché p^2 è zero a $p = 0$ e $e^{2\pi p}$ tende a zero quando p tende a $-\infty$. Quindi, ci rimane solo il secondo termine, che è un altro integrale che può essere risolto applicando nuovamente l'integrazione per parti

$$\int_{-\infty}^{0} p e^{2\pi p} \mathrm{d}p = \cancel{p \frac{1}{2\pi} e^{2\pi p}\Big|_{-\infty}^{0}} - \int_{-\infty}^{0} \frac{1}{2\pi} e^{2\pi p} \mathrm{d}p.$$

Anche in questo caso, il primo termine si annulla perché p è zero in $p = 0$ ed $e^{2\pi p}$ tende a zero quando p tende a $-\infty$. Quindi, rimane solo il secondo termine, che può essere immediatamente calcolato come

$$\int_{-\infty}^{0} e^{2\pi p} \mathrm{d}p = \frac{1}{2\pi} e^{2\pi p}\Big|_{-\infty}^{0} = \frac{1}{2\pi}.$$

Infine, mettendo insieme tutto, otteniamo che

$$\|f(x)\|_2 = \frac{3\sqrt{2\pi}}{2\pi^3}. \tag{4.19}$$

[a] La funzione $u(x)$ è nota come *funzione gradino di Heaviside*.

Esempio 4.2. Si calcoli la trasformata di Fourier della funzione $f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$. Allora

$$\begin{aligned}\mathcal{F}[f] &= \frac{1}{\pi\sqrt{2\pi}}\int_{\mathbb{R}} \frac{\mathrm{e}^{-ipx}}{1+x^2}\,\mathrm{d}x = \frac{1}{\pi\sqrt{2\pi}}\int_{\mathbb{R}} \frac{\mathrm{e}^{-ipz}}{1+z^2}\,\mathrm{d}z \\ &= \begin{cases} \frac{1}{\pi\sqrt{2\pi}} 2\pi i \,\mathrm{Res}\{f,i\} & p<0 \\ \frac{1}{\pi\sqrt{2\pi}} 2\pi i \,\mathrm{Res}\{f,-i\} & p>0 \end{cases} = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-|p|}\end{aligned} \tag{4.20}$$

Non è difficile verificare che antitrasformando la (4.20) si riottiene la $f(x)$ di partenza.

Esempio 4.3. Si calcoli la trasformata della funzione $g(x) = \mathbb{1}_{[-\frac{T}{2},\frac{T}{2}]}(x)$. Si ha che

$$\mathcal{F}[g] = \frac{1}{\sqrt{2\pi}}\int_{-T/2}^{T/2} \mathrm{e}^{-ipx}\,\mathrm{d}x = \sqrt{\frac{2}{\pi}}\frac{\sin\left(\frac{T}{2}p\right)}{p} \tag{4.21}$$

Per l'antitrasformata, si devono fare considerazioni un po' più accurate. Infatti, scrivendo esplicitamente la formula

$$\begin{aligned}\mathcal{F}^{-1}[\hat{g}] &= \frac{1}{\pi}\int_{\mathbb{R}} \frac{\sin\left(\frac{T}{2}p\right)}{p}\mathrm{e}^{-ipx}\,\mathrm{d}p = \frac{1}{\pi}\int_{\mathbb{R}} \frac{1}{p}\frac{\mathrm{e}^{i\frac{T}{2}p} - \mathrm{e}^{-i\frac{T}{2}p}}{2i}\mathrm{e}^{ipx}\,\mathrm{d}x \\ &= \frac{1}{2\pi i}\int_{\mathbb{R}} \left[\frac{\mathrm{e}^{ip(\frac{T}{2}+x)}}{p} - \frac{\mathrm{e}^{ip(x-\frac{T}{2})}}{p}\right]\mathrm{d}p\end{aligned} \tag{4.22}$$

Si sarebbe tentati di spezzare l'integrale sfruttandone la linearità; così facendo però si creerebbero divergenze fittizie in $p = 0$ che comprometterebbero la bontà del calcolo. Si sfrutta allora il risultato secondo cui l'integrale di una funzione analitica non cambia se la funzione viene integrata su una curva **omotopa**[a] a quella di partenza. Si definisce Γ_+ come la curva in cui si sostituisce all'intervallo $[-R,+R]$ la semicirconferenza centrata in $p = 0$ di raggio R nel semipiano delle $\Im(z) > 0$. A questo punto, non essendoci più singolarità lungo il cammino, l'integrale può essere spezzato e si può usare il teorema dei residui, facendo tendere $R \to \infty$. Calcolando opportunamente i due termini, si riottiene $g(x)$. Si noti che la scelta di Γ_+ è solo una delle scelte possibili, dettata da un criterio di semplicità. Come poc'anzi affermato, qualunque curva omotopa ad $\mathbb{R}$ andrebbe bene per fare questo calcolo.

[a] In generale, due curve sono dette omotope se una delle due può essere *deformata con continuità* nell'altra. Tale trasformazione è detta **omotopia** tra le due curve.

4.2.3 Applicazioni alle equazioni differenziali

La trasformata di Fourier può essere uno strumento molto utile per risolvere alcuni tipi di equazioni differenziali. Un esempio è[2] il seguente.

Esempio 4.4. Cerchiamo una soluzione particolare di[a]

$$-y'' + y = 1 \tag{4.23}$$

[2] Fabio Bagarello (✝★1964), *Metodi matematici per fisici e ingegneri*, Zanichelli, Seconda Edizione, 2019.

Osserviamo che $y_p(x) \equiv 1$ risolve banalmente l'equazione. Applichiamo ad ambo i membri l'operatore di trasformazione di Fourier, sfruttandone la linearità. Si ha quindi, ricordando la proprietà (1) di 4.1 che

$$\mathcal{F}(-y'') + \mathcal{F}(y) = \mathcal{F}(1) \Rightarrow (p^2+1)\hat{f}(p) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ipx} dx = \sqrt{2\pi}\delta(p)$$

$$\Longrightarrow \hat{f}(p) = \frac{\sqrt{2\pi}}{p^2+1}\delta(p) = \sqrt{2\pi}\delta(p),$$

la cui antitrasformata è proprio $y_p(x)$.

[a] In questo esempio si fa uso della delta di Dirac, di cui si parlerà in seguito.

4.3 Trasformata di Laplace

Tra le trasformate integrali, di particolare importanza per la risoluzione di alcune classi di equazioni differenziali (e non solo) sono le *trasformate di Laplace*. A differenza di una trasformata di Fourier, si mostrerà che il risultato di una trasformazione di Laplace è una funzione complessa analitica. Questo è particolarmente utile perché permette di associare funzioni estremamente regolari[3] a funzioni che regolari possono anche esserlo poco.

Definizione 4.5. *Data una funzione reale $f(t)$, la sua* **trasformata di Laplace** *(unilaterale) è la funzione $F(z)$ complessa di variabile complessa, ottenuta dalla formula*

$$F(z) = \mathcal{L}\left[f(t)\right](z) = \int_0^{+\infty} e^{-zt} f(t)\, dt \tag{4.24}$$

Per studiare il dominio della trasformata, si introduce lo spazio delle funzioni *localmente* sommabili o integrabili.

Definizione 4.6. *Lo spazio*

$$\mathscr{L}^1_{loc}(\mathbb{R}) = \left\{ f(t) \;\; \mathscr{L}\text{-misurabile} \quad \text{t.c.} \quad \forall \mathcal{E} \subset \mathbb{R} : \mu(\mathcal{E}) < \infty \Rightarrow \int_{\mathcal{E}} |f(t)|\, dt < \infty \right\}$$

è detto delle funzioni **localmente integrabili**.

Non è difficile mostrare che vale la

Proposizione 4.9. $\mathscr{L}^p(\mathbb{R}) \subset \mathscr{L}^1_{loc}(\mathbb{R}) \quad \forall\, 1 \leq p \leq +\infty$

Dimostrazione. Sia $f \in \mathscr{L}^p$ con p fissato. Allora $\|f\|_p = \left(\int_{\mathbb{R}} |f|^p dx\right)^{1/p} < +\infty$. Allora, preso $\mathcal{E} \subset \mathbb{R}$ a misura finita e utilizzando la disuguaglianza di Hölder[a] per il prodotto $f\mathbb{1}_{\mathcal{E}}$, si ha $\int_{\mathcal{E}} |f| dx = \int_{\mathbb{R}} |f\mathbb{1}_{\mathcal{E}}| dx = \|f\mathbb{1}_{\mathcal{E}}\|_1 \leq \|f\|_p \|\mathbb{1}_{\mathcal{E}}\|_q = (\mu(\mathcal{E}))^{1/q} \|f\|_p < +\infty$, ove q è l'esponente coniugato a p, cfr. nota. Dall'arbitrarietà di $\mathcal{E}$ e di p, segue la tesi. □

[3] Le funzioni analitiche sono di classe C^∞.

[a]La **disuguaglianza di Hölder** afferma che, dati p, q esponenti coniugati, ovvero tali che $\frac{1}{p} + \frac{1}{q} = 1$, se $f \in \mathscr{L}^p$ e $g \in \mathscr{L}^q$ allora $fg \in \mathscr{L}^1$ e vale inoltre che $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

Non vale in generale l'uguaglianza per alcun valore di p, cioè esistono funzioni di $\mathscr{L}^1_{loc}(\mathbb{R})$ che non stanno in $\mathscr{L}^p$, ad esempio $f(t) = e^t$. Si possono a questo punto studiare le condizioni affinché l'integrale in (4.24) converga.

Definizione 4.7. *Siano $M > 0, \alpha \in \mathbb{R}, t_0 \geq 0$, lo spazio*

$$\mathcal{F}_{M,\alpha,t_0} = \left\{ f(t) \in \mathscr{L}^1_{loc}(\mathbb{R}) : |f(t)| \leq M e^{\alpha t} \; q.o. \forall t > t_0 \right\} \tag{4.25}$$

è detto delle funzioni che divergono in modulo più lentamente di ogni esponenziale lineare.

È ovvio che $\mathcal{F}_{M,\alpha,t_0} \subset \mathscr{L}^1_{loc}(\mathbb{R})$, in quanto ad esempio la funzione $f(t) = e^{t^2}$ sta in $\mathscr{L}^1_{loc}(\mathbb{R})$ ma non in $\mathcal{F}_{M,\alpha,t_0}$ quali che siano i parametri scelti.

Non è difficile convincersi della seguente proprietà.

Proposizione 4.10. *Se $f \in \mathcal{F}_{M,\alpha,t_0}$, allora $f \in \mathcal{F}_{M',\alpha',t'_0}$, per ogni $M' \geq M, \alpha' \geq \alpha, t'_0 \geq t_0$.*

In generale non vale il viceversa, nel senso che l'inclusione inversa può valere in alcuni casi, ma non si può concludere nulla di generale. Gli spazi $\mathcal{F}$ sono importanti in quanto costituiscono un dominio naturale per la trasformata di Laplace.

Teorema 4.4. *Se $f(t) \in \mathcal{F}_{M,\alpha,t_0}$, allora $\exists F(z) = \mathcal{L}[f(t)] = \int_0^{+\infty} e^{-zt} f(t)\, dt \;\; \forall z \in \mathbb{C} : \Re\{z\} > \alpha$.*

Dimostrazione. Detto $z = x + iy$, allora

$$\left| \int_0^{+\infty} f(t) e^{-zt}\, dt \right| \leq \int_0^{+\infty} |f(t)| e^{-xt}\, dt \leq (\star)$$

in quanto $|e^{-zt}| = |e^{-xt}| \cancel{|e^{-iyt}|} = e^{-xt}$. Inoltre,

$$(\star) = \int_0^{t_0} |f(t)| e^{-xt}\, dt + \int_{t_0}^{+\infty} |f(t)| e^{-xt}\, dt \leq \sup_{t \in [0,t_0]} (e^{-xt}) \underbrace{\int_0^{t_0} |f(t)|\, dt}_{<+\infty} + M \int_{t_0}^{+\infty} e^{(\alpha - x)t}\, dt < \infty$$

se $x > \alpha$. Questo dimostra la tesi. □

Questo teorema mostra che se la trasformata è definita per un certo valore di $\Re(z)$, allora è definita per tutti gli $z \in \mathbb{C}$ aventi parte reale maggiore o uguale di tale valore. Non dà però informazioni sulle z complesse avente parte reale minore: nulla esclude che anche per quei valori possa esistere la trasformata.

In generale, $\exists! \,\hat{\alpha}_f := \inf\{\alpha : f(t) \in \mathcal{F}_{M,\alpha,t_0}\}$, eventualmente anche uguale a $-\infty$. Tale valore ha il seguente significato: $\forall z \in \mathbb{C}$ con $\Re(z) > \hat{\alpha}_f$ la trasformata di Laplace esiste sicuramente, mentre per $\Re(z) < \hat{\alpha}_f$ la trasformata non è definita. Per questo motivo, si può dare la seguente definizione.

Definizione 4.8. *Il valore* $\hat{\alpha}_f := \inf\{\alpha : f(t) \in \mathcal{F}_{M,\alpha,t_0}\}$ *prende il nome di* **ascissa di convergenza** *della trasformata di Laplace della funzione* $f(t)$.

In generale, nulla si può dire quando $\Re(z) = \hat{\alpha}_f$ e saranno possibili punti su tale retta in cui la trasformata è definita e punti in cui essa non è definita. Nel caso delle funzioni continue a supporto compatto, si può dimostrare (esercizio per il lettore!) che $\hat{\alpha}_f = -\infty$. L'ascissa di convergenza si può anche determinare analizzando $f(t)$ oppure a posteriori, dopo aver trasformato la funzione e capendo dove ha senso definirla.

Ad esempio, si consideri la funzione $f(t) = t$. Essa è tale che $t < e^{\epsilon t}$ $\forall \epsilon > 0$. Allora $f(t) \in \mathcal{F}_{M,\epsilon,t_0}$ $\forall \epsilon > 0$ e quindi l'ascissa di convergenza è $\hat{\alpha}_f = 0$. Si può ora dimostrare il seguente importante teorema.

Teorema 4.5. *Se* $f(t) \in \mathcal{F}_{M,\alpha,t_0}$, *allora* $F(z)$ *è analitica.*

Dimostrazione. Si calcoli la derivata prima e si mostri che è ben definita. Risulta che $F'(z) = \int_0^{+\infty} \frac{\partial}{\partial z} \left(e^{-zt} f(t)\right) dt = -\int_0^{+\infty} f(t) t e^{-zt}\, dt = -\mathcal{L}\left[t f(t)\right](z)$. Infatti, la funzione $g(t) = t f(t)$ ammette trasformata di Laplace in quanto $|t f(t)| \leq |t||f(t)| \leq e^{\epsilon} M e^{\hat{\alpha}_f t} = M e^{(\hat{\alpha}_f + \epsilon)t}$ e quindi g e f hanno la stessa ascissa di convergenza. Dunque F è derivabile $\Rightarrow$ è olomorfa $\Rightarrow$ è analitica. □

Da questa dimostrazione si ricava la seguente proprietà.

Proposizione 4.11. $F^{(n)}(z) = (-1)^n \mathcal{L}\left[t^n f(t)\right](z)$ *e ha la stessa ascissa di convergenza di* $F(z)$.

4.3.1 Antitrasformata di Laplace

La trasformata di Laplace (unilaterale) è definita mediante un'integrazione su $[0, +\infty]$. Se quindi due funzioni differiscono soltanto per $t < 0$ ma sono uguali q.o. per $t > 0$, è evidente che avranno la stessa trasformata. Questo fatto suggerisce che l'antitrasformata di una funzione analitica è definita a meno di un insieme di misura nulla, ovvero restituisce la funzione di partenza soltanto q.o. Inoltre, poiché nel calcolo della trasformata entra soltanto il comportamento della funzione $f(t)$ nell'intervallo $[0, +\infty]$, è del tutto naturale aspettarsi che il risultato dell'antitrasformata restituisca la funzione in $[0, +\infty]$ e 0 altrove.

Queste considerazioni del tutto euristiche devono venir fuori matematicamente quando si definisce l'antitraformata. Per ricavare l'antitrasformata di Laplace, ci si riconduce a studiare una trasformata di Fourier riscrivendo (4.24) come

$$\mathcal{L}\left[f(t)\right] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left[u(t) f(t) \sqrt{2\pi} e^{-xt}\right] e^{-iyt}\, dt = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \Phi_x(t) e^{-iyt}\, dt = \mathcal{F}\left[\Phi_x\right] \tag{4.26}$$

dove $u(t) = \mathbb{1}_{\mathbb{R}^+}$ è la funzione di Heaviside e vale 1 per t positivo e 0 altrove. Da questo si ricava che

$$\Phi_x(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} F(z) e^{-iyt} dy \tag{4.27}$$

$$\Rightarrow u(t) f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} F(x + iy) e^{(x+iy)t}\, dt = \frac{1}{2\pi i} \int_{x-i\infty}^{x+i\infty} F(z) e^{zt}\, dz \tag{4.28}$$

Si può quindi porre la seguente definizione.

Definizione 4.9. *Data una funzione analitica $F(z)$, la sua* **antitrasformata di Laplace** *si calcola mediante la legge*

$$\mathcal{L}^{-1}[F(z)](t) = \frac{1}{2\pi i} \int_{x-i\infty}^{x+i\infty} F(z) e^{zt} \, dz \tag{4.29}$$

dove l'integrazione viene eseguita lungo una qualsiasi retta, parallela all'asse y, con $x > \hat{\alpha}_f$ (ascissa di convergenza).

Proposizione 4.12. *La legge* (4.29) *restituisce la funzione $f(t)$ q.o. per $t > 0$, ovvero restituisce $u(t)f(t)$.*

4.3.2 Proprietà di $\mathcal{L}$ e di $\mathcal{L}^{-1}$

Riassiumiamo qui alcune proprietà utili nello svolgimento degli esercizi.

1. $\mathcal{L}$ e $\mathcal{L}^{-1}$ sono operatori lineari.
2. Vale la 4.11, ovvero
$$\mathcal{L}\left[t^n f(t)\right] = (-1)^n \frac{d^n}{dz^n} \mathcal{L}\left[f(t)\right](z). \tag{4.30}$$
3. Valgono le seguenti relazioni (che invitiamo a dimostrare)
$$\mathcal{L}\left[e^{z_0 t} f(t)\right](z) = \mathcal{L}\left[f(t)\right](z - z_0), \tag{4.31}$$
$$\mathcal{L}^{-1}\left[e^{-az} F(z)\right] = u(t-a) f(t-a), \tag{4.32}$$
$$\mathcal{L}\left[f(t - t_0)\right] = e^{-zt_0} \mathcal{L}\left[F(t)\right], \tag{4.33}$$
$$\mathcal{L}\left[f(at)\right](z) = \frac{1}{a} \mathcal{L}\left[f(t)\right]\left(\frac{z}{a}\right) \quad \forall a > 0. \tag{4.34}$$

Seguono alcuni esempi di calcolo della trasformata di Laplace.

Esempio 4.5. Si consideri la funzione $f(t) = 1$. È immediato calcolare la sua trasformata:

$$F(z) = \mathcal{L}\left[f(t)\right](z) = \int_0^{+\infty} e^{-zt} \, dt = \frac{1}{z}. \tag{4.35}$$

L'ascissa di convergenza è evidentemente $\alpha = 0$. Per il calcolo dell'antitrasformata può essere utile una versione modificata del **lemma di Jordan**, cfr. A.1. Si sta infatti integrando lungo una retta parallela all'asse y. Allora, con riferimento al lemma di Jordan A.2, si è sicuri che l'integrale sulla semicirconferenza contenuta nel semipiano alla destra della retta tenda a 0 al divergere di R solo se $t < 0$; viceversa, l'integrale sulla semicirconferenza contenuta nel semipiano alla sinistra della retta tende a 0 al divergere di R se $t > 0$. Quindi, scelto $x > 0$:

$$\mathcal{L}^{-1}[F(z)](t) = \lim_{R,+\infty} \frac{1}{2\pi i} \int_{x-iR}^{x+iR} \frac{e^{zt}}{z} \, dz = \begin{cases} 0 & t < 0 \\ \text{Res}\{f, 0\} & t > 0 \end{cases} = \begin{cases} 0 & t < 0 \\ 1 & t > 0 \end{cases} = u(t) f(t). \tag{4.36}$$

Conoscendo la trasformata di Laplace di tale funzione, e sfruttando la Proposizione 4.11, si ottiene immediatamente che $\mathcal{L}\left[t^n\right] = \frac{n!}{z^{n+1}}$.

Esempio 4.6. Si consideri ora la funzione $g(t) = \mathbb{1}_{[0,1]}(t)$. Non è difficile calcolare che $F(z) = \frac{1-\mathrm{e}^{-z}}{z}$, con ascissa di convergenza pari ad $\alpha = -\infty$. Questo è un caso particolare di una situazione del tutto generale: ogni trasformata di funzioni continue a supporto compatto ha ascissa di convergenza pari a $-\infty$.

4.3.3 Applicazioni alle equazioni differenziali

Per la trasformata di Laplace della derivata di una funzione vale la seguente legge, dimostrabile per induzione.

Proposizione 4.13 (Eq. diff.). *Data la funzione f derivabile n volte, si ha che*

$$\mathcal{L}\left[\frac{\mathrm{d}^n f}{\mathrm{d}t^n}\right] = z^n \mathcal{L}\left[f\right](z) - z^{n-1} f(0) - z^{n-2} f'(0) - \cdots - f^{(n-1)}(0) \tag{4.37}$$

Questa proprietà, che di fatto trasforma a meno di un termine derivate in polinomi, può essere utilizzata per risolvere equazioni differenziali lineari a coefficienti costanti. Sia infatti

$$\sum_{k=0}^{n} c_k \frac{\mathrm{d}^k y}{\mathrm{d}t^k} = f(t) \tag{4.38}$$

un'equazione lineare di ordine n e si consideri il problema di Cauchy dotato dell'insieme di condizioni iniziali $\{y(0) = y_0, y'(0) = y_1, \ldots, y^{(n-1)}(0) = y_{n-1}\}$. Si applichi ora ad ambo i membri di (4.38) l'operatore di Laplace. Questo è possibile e mantiene l'uguaglianza in quanto se due funzioni sono uguali, avranno anche stessa trasformata di Laplace. Considerando che, secondo (4.37), la trasformata di una derivata può essere scritta come la somma di un termine di ordine k avente per coefficiente $\mathcal{L}[y]$ e di un polinomio di ordine $k-1$ avente coefficienti determinati dalle condizioni iniziali, si ottiene infine

$$Y(z)(c_0 + c_1 z + \cdots + c_n z^n) + P_{n-1}(z) = F(z) \tag{4.39}$$

dove si è posto $Y(z) = \mathcal{L}[y]$, $F(z) = \mathcal{L}[f]$ e $P_{n-1}(x)$ è un polinomio di ordine $n-1$ i cui coefficienti sono un'opportuna combinazione delle condizioni iniziali. A questo punto, si ricava che

$$\mathcal{L}[y](z) = Y(z) = \frac{F(z) - P_{n-1}(z)}{c_0 + c_1 z + \cdots + c_n z^n} \tag{4.40}$$

e, antitrasformando, si ottiene la soluzione dell'equazione differenziale (4.38) di partenza per $t > 0$. Per sua stessa natura, infatti, l'operazione di antitrasformazione restituisce $u(t)y(t)$, cancellando ogni informazione sulla soluzione per $t < 0$. È importante notare come il risultato di questa procedura fornisca la soluzione che soddisfa le condizioni iniziali del problema, a differenza di altri metodi di risoluzione in cui si ottiene una soluzione generale definita a meno di n parametri arbitrari che devono essere fissati in seguito risolvendo sistemi di equazioni.

Esempio 4.7. Si consideri il problema di Cauchy:

$$\begin{cases} y' + y = \mathrm{e}^{-t} \\ y(0) = 1 \end{cases} \tag{4.41}$$

Si ha che $\mathcal{L}\left[\mathrm{e}^{-t}\right] = \frac{1}{1+z}$, con ascissa di convergenza $\alpha = -1$. Invece $\mathcal{L}[y' + y] = \mathcal{L}[y] + z\mathcal{L}[y] - y(0) = Y(z)(1+z) - 1$. Allora, la trasformata di Laplace della soluzione è $Y(z) = \frac{2+z}{(1+z)^2}$ e,

antitrasformando

$$u(t)y(t) = \frac{1}{2\pi i} \int_{x-i\infty}^{x+i\infty} \frac{2+z}{(1+z)^2} e^{zt}\, dz = \text{Res} \left\{ \frac{2+z}{(1+z)^2} e^{zt} \right\} \bigg|_{z=-1} = (1+t)e^{-t}. \tag{4.42}$$

Capitolo 5

Cenni sulle distribuzioni

5.1 Introduzione

Le *distribuzioni* possono essere viste come una generalizzazione del concetto di funzione. Storicamente, la loro introduzione fu motivata da alcuni problemi di natura formale derivanti dall'utilizzo della *delta di Dirac*, importante per il suo utilizzo in meccanica quantistica e in elettrodinamica.

La $\delta(x)$ viene *definita* dalla sua azione su una funzione sotto il segno di integrale. Data una funzione $f \in \mathscr{L}^2(\mathbb{R})$, la $\delta(x)$ è quell'oggetto tale che

$$\int_{\mathbb{R}} f(x)\delta(x - x_0)\,\mathrm{d}x = f(x_0) \tag{5.1}$$

ovvero, che restituisce la funzione in un punto x_0. Una rappresentazione integrale di tale oggetto si ottiene nel linguaggio delle trasformate di Fourier.

Definizione 5.1 (Rappresentazione integrale della δ di Dirac). *La delta di Dirac $\delta(x)$ si può rappresentare come*

$$\delta(x) := \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{ipx}\,\mathrm{d}p\,. \tag{5.2}$$

Questa uguaglianza *debole* si ottiene, per l'appunto, da una trasformata di Fourier. Ricordando che

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \hat{f}(x)\,\mathrm{e}^{ipx}\mathrm{d}p \quad \text{e} \quad \hat{f}(y) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(y)\,\mathrm{e}^{-ipy}\mathrm{d}y$$

e sostituendo la seconda nella prima, si ottiene, per l'appunto

$$\begin{aligned} f(x) &= \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{ipx}\left[\int_{\mathbb{R}} \mathrm{e}^{-ipy} f(x)\mathrm{d}y\right]\mathrm{d}p = \int_{\mathbb{R}}\left[\frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{ipx}\mathrm{e}^{-ipy}\mathrm{d}p\right] f(y)\mathrm{d}y \\ &= \int_{\mathbb{R}}\left[\frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{ip(x-y)}\mathrm{d}p\right] f(y)\mathrm{d}y \equiv \int_{\mathbb{R}} \mathrm{d}y\,\delta(x-y)f(y), \end{aligned} \tag{5.3}$$

da cui risulta evidente che, affinché l'uguaglianza sussista, l'integrale entro parentesi quadre debba rappresentare $\delta(x - y)$. Non è difficile rendersi conto che l'integrale definito in (5.2) diverge per $x = 0$, mentre non esiste per $x \neq 0$.

Queste definizioni, apparentemente innocue, pongono problemi non indifferenti. Si supponga per assurdo che la $\delta(x)$ sia una funzione, ad esempio una funzione di $\mathscr{L}^1_{loc}(\mathbb{R})$: allora, scelte le tre funzioni

$$f(x) \in \mathscr{L}^2(\mathbb{R}), \qquad g(x) = \begin{cases} f(x) & x \neq x_0 \\ f(x_0) + \sin(x) & x = x_0 \end{cases}, \qquad h(x) = \begin{cases} f(x) & x \neq x_0 \\ f(x_0) - 10^{18} & x = x_0 \end{cases}$$

si ha $f \overset{L}{\equiv} g \overset{L}{\equiv} h$, ovvero hanno lo stesso integrale di Lebesgue. Tuttavia, $\int_{\mathbb{R}} f(x)\,\delta(x-x_0)\,\mathrm{d}x \neq \int_{\mathbb{R}} g(x)\,\delta(x-x_0)\,\mathrm{d}x \neq \int_{\mathbb{R}} h(x)\,\delta(x-x_0)\,\mathrm{d}x$, contrariamente a quanto previsto dal teorema B.1. Il problema sorge non appena si considerano funzioni $\mathscr{L}$-equivalenti tra loro, senza nulla specificare circa la loro regolarità.

Un modo per superare il problema potrebbe essere restringersi a considerare funzioni di $\mathscr{S}(\mathbb{R}) \subset \mathscr{L}^2(\mathbb{R})$ o in $\mathscr{L}^2(\mathbb{R}) \cap C(\mathbb{R})$. Nel primo caso, in questo spazio due funzioni sono $\mathscr{L}$-equivalenti se, e solo se, coincidono su tutto $\mathbb{R}$, nel secondo caso le funzioni sono anche continue.

Questo, però, non risolve il problema. Si può infatti dimostrare che

Teorema 5.1 (*No go*, v. 1). **Non esiste** *alcuna funzione non negativa $t(x) \in \mathscr{L}^1_{loc}(\mathbb{R})$ tale da soddisfare $\int_{\mathbb{R}} t(x)f(x)\mathrm{d}x = f(0) \;\; \forall\, f(x) \in \mathscr{S}(\mathbb{R})$.*

Dimostrazione. Si supponga per assurdo che tale funzione esista. Si avrebbe che $t(x)$ è una funzione $\mathscr{L}^1_{loc}(\mathbb{R})$. Se la proprietà vale $\forall f \in \mathscr{S}(\mathbb{R})$, in particolare vale per $f(x) = x^2 \mathrm{e}^{-x^2}$. Tale funzione è non negativa in tutto il dominio e nulla per $x = 0$. Di conseguenza, vale $\int_{\mathbb{R}} f(x)\,t(x)\,\mathrm{d}x = f(0) = 0$. Per quanto già detto, l'integrale può essere nullo se, e solo se, $f(x)\,t(x) = 0$ q.o., ovvero se, e solo se $t(x) = 0$ q.o. Sia ora $f(x) = \mathrm{e}^{-x^2}$. Si ha quindi che

$$0 = \int_{\mathbb{R}} f(x)t(x)\,\mathrm{d}x = f(0) = 1 \quad \lightning$$

che porta, per l'appunto, a un assurdo. □

Le ipotesi di questo teorema possono essere alleggerite; in seguito si dimostrerà che non esiste alcuna funzione almeno localmente integrabile che abbia le proprietà di una delta.

Per introdurre il concetto di distribuzione, è necessario affrontare prima alcuni concetti ausiliari. Ricordiamo la definizione dello spazio delle funzioni a decrescenza rapida.

Definizione 5.2. *Lo spazio delle funzioni a decrescenza rapida è definito come*

$$\mathscr{S}(\mathbb{R}) := \left\{ f(x) \in C^\infty \mid \lim_{|x|,+\infty} |x|^k f^{(l)}(x) \; \forall k,l \in \mathbb{N} \cup \{0\} \right\} \tag{5.4}$$

ovvero lo spazio di tutte quelle funzioni continue, infinitamente derivabili, e che vanno a zero, insieme a tutte le loro derivate, più velocemente di qualsiasi monomio.

Come è già noto, si ha che

Proprietà 5.1. *$\mathscr{S}(\mathbb{R})$ è denso in $\mathscr{L}^2(\mathbb{R})$, ovvero*

$$f(x) \in \mathscr{L}^2(\mathbb{R}) \implies \forall \epsilon > 0\, \exists f_0(x) \in \mathscr{S}(\mathbb{R}) : \|f - f_0\| < \epsilon.$$

Poiché $\mathscr{S}(\mathbb{R})$ è denso in $\mathscr{L}^2(\mathbb{R})$, sarà chiaramente possibile trovare una successione di funzioni di $\mathscr{S}(\mathbb{R})$ che non converge in tale spazio. Questo fatto è riassunto nella seguente

Proprietà 5.2. *$\mathscr{S}(\mathbb{R})$ non è completo, ovvero, data una successione $\{f_n(x)\}_{n\in\mathbb{N}} \in \mathscr{S}$ tale che $\forall\, n, m > n_\epsilon$ $\|f_n(x) - f_m(x)\| < \epsilon$, si ha che*

$$\exists\, f(x) = \|\cdot\|_2\text{-}\lim_{n\to\infty} f_n \in \mathscr{L}^2(\mathbb{R}).$$

Pertanto, si può approssimare "bene a piacere" una funzione di $\mathscr{L}^2(\mathbb{R})$ rimanendo in $\mathscr{S}(\mathbb{R})$. Un'altra proprietà che tornerà utile nelle applicazioni è la seguente

Proprietà 5.3. *Date $f(x) \in \mathscr{S}(\mathbb{R})$ e $\varphi(x) \in \mathscr{L}^\infty(\mathbb{R}) \cap C^\infty(\mathbb{R})$ si ha che $g(x) := f\varphi \in \mathscr{S}(\mathbb{R})$.*

Per introdurre le distribuzioni, è utile definire un ulteriore spazio funzionale. Al fine di ciò, introduciamo il concetto di

Definizione 5.3 (supporto). *Il* supporto *di una funzione è l'insieme definito come*

$$\text{supp}\{f(x)\} := \overline{\{x \in \text{dom}(f) \,:\, f(x) \neq 0\}} \tag{5.5}$$

ovvero la chiusura dell'insieme dei punti del dominio di f in cui quest'ultima non si annulla.

Adesso, possiamo introdurre lo spazio di Dirichlet.

Definizione 5.4. *Si definisce $\mathcal{D}(\mathbb{R})$ come lo spazio*

$$\mathcal{D}(\mathbb{R}) := \{f(x) \in C^\infty : \text{supp}\{f\} \textit{ sia compatto}\} \tag{5.6}$$

ovvero delle funzioni reali di variabile reale di classe C^∞ e a supporto compatto, ovvero chiuso e limitato.

Esempio 5.1. Un esempio di funzione appartenente a $\mathcal{D}(\mathbb{R})$ è

$$h(x) = \begin{cases} e^{\frac{1}{x^2-1}} & x \in]\text{-}1,1[\\ 0 & \text{altrove} \end{cases} \tag{5.7}$$

Non è difficile infatti verificare che la funzione è continua. In generale, invece, la derivata n-esima di $h(x)$ può essere scritta come

$$h^{(n)}(x) = \begin{cases} R(x)\, e^{\frac{1}{x^2-1}} & x \in]\text{-}1,1[\\ 0 & \text{altrove} \end{cases} \tag{5.8}$$

dove $R(x)$ è una funzione razionale fratta avente delle singolarità solo in $x = \pm 1$. È evidente che $\lim_{|x|,1} h^{(n)}(x) = 0$ in quanto l'esponenziale è un infinitesimo di ordine superiore rispetto a qualsiasi polinomio. Visto che $h(x) \in \mathcal{D}(\mathbb{R})$, sicuramente staranno nello stesso spazio anche le funzioni

$$h_a(x) = \begin{cases} \exp\left[(x^2 - a^2)^{-1}\right] & x \in] - a, a[\\ 0 & \text{altrove} \end{cases} \tag{5.9}$$

e, in generale, anche le funzioni[a] nella forma $f(x) = p(x)\, h_a(x)$, dove $p(x)$ è una funzione qualsiasi di classe C^∞.

[a] Non è detto però che ogni funzione di $\mathcal{D}(\mathbb{R})$ sia espandibile in questa forma.

Non è difficile concludere che $\mathcal{D}(\mathbb{R}) \subset \mathscr{S}(\mathbb{R}) \subset \mathscr{L}^2(\mathbb{R})$, in quanto una funzione di $\mathcal{D}(\mathbb{R})$ è di classe C^∞ e all'infinito è esattamente uguale a zero con tutte le sue derivate. Quindi, anche $\mathcal{D}(\mathbb{R})$

è denso in $\mathscr{L}^2(\mathbb{R})$. La densità di $\mathcal{D}(\mathbb{R})$ e di $\mathscr{S}(\mathbb{R})$ in $\mathscr{L}^2(\mathbb{R})$ mostrano come i due spazi non siano completi rispetto alla $\|\cdot\|_2$. In generale:

Teorema 5.2. *$\mathcal{D}(\mathbb{R})$ e $\mathscr{S}(\mathbb{R})$ sono densi in $\mathscr{L}^p(\mathbb{R})$ nella $\|\cdot\|_p$, ovvero in simboli*

$$\overline{\mathcal{D}(\mathbb{R})}^{\|\cdot\|_p} = \overline{\mathscr{S}(\mathbb{R})}^{\|\cdot\|_p} = \mathscr{L}^p(\mathbb{R}) \ , \ 1 \leq p \leq +\infty.$$

Nonostante questo, in entrambi i casi è possibile definire delle topologie che rendono tali spazi completi.

Definizione 5.5. *Una successione $\{\varphi_j\}$ di funzioni di $\mathcal{D}(\mathbb{R})$ converge nella topologia $\tau_{\mathcal{D}}$ a una φ, ovvero esiste il $\tau_{\mathcal{D}}\text{-}\lim\limits_{j,+\infty} \varphi_j = \varphi$, se*

1. *$\exists\, I \subset \mathbb{R}$ limitato tale che $\operatorname{supp}\{\varphi_j\} \subseteq I\ \forall j$ e $\operatorname{supp}\{\varphi\} \subseteq I$;*

2. *$\varphi_j^{(k)} \rightrightarrows \varphi^{(k)}$ (converge uniformemente) $\forall k \in \mathbb{N} \cup \{0\}$;*

Teorema 5.3. *La topologia $\tau_{\mathcal{D}}$ rende lo spazio $\mathcal{D}(\mathbb{R})$ completo.*

Proposizione 5.1. *Data una successione $\{\varphi_j\} \subset \mathscr{S}(\mathbb{R})$, la topologia $\tau_{\mathscr{S}}$ nella quale il*

$$\tau_{\mathscr{S}}\text{-}\lim_{j,+\infty} \varphi_j = \varphi$$

è quella in cui si ha $x^k\, \varphi_j^{(l)} \rightrightarrows x^k\, \varphi^{(l)}$ per ogni $k, l \in \mathbb{N} \cup \{0\}$, ovvero equivalentemente

$$\lim_{j,+\infty} \sup_{x \in \mathbb{R}} \left| x^k \left[\varphi_j^{(l)}(x) - \varphi^{(l)}(x) \right] \right| = 0. \tag{5.10}$$

Tale topologia rende lo spazio $\mathscr{S}(\mathbb{R})$ completo.

A questo punto possiamo dimostrare il Teorema 5.1 in termini più generali.

Teorema 5.4 (*No go, v. 2*). **Non esiste** *alcuna funzione $t(x) \in \mathscr{L}^1_{loc}(\mathbb{R})$ tale da soddisfare $\int_{\mathbb{R}} t(x) f(x) \mathrm{d}x = f(0)\ \forall f(x) \in \mathscr{S}(\mathbb{R})$*

Dimostrazione. Riprendiamo la definizione di $h_1(nx)$. Essa è definita come

$$h_1(nx) = \begin{cases} \exp\{(n^2x^2 - 1)^{-1}\} & |x| < 1/n \\ 0 & \text{altrove} \end{cases}$$

Sostituiamo quindi $h_1(nx)$ a $f(x)$ con $n \in \mathbb{N}$. Di conseguenza, si ha:

$$\int_{\mathbb{R}} t(x) h_1(nx) \mathrm{d}x = h(0) = \mathrm{e}^{-1} = \int_{-1/n}^{1/n} t(x) \mathrm{e}^{\frac{1}{n^2x^2-1}} \mathrm{d}x$$

che, passando ai valori assoluti, diventa

$$\left| \frac{1}{\mathrm{e}} \right| = \left| \int_{-1/n}^{1/n} t(x) \mathrm{e}^{\frac{1}{n^2x^2-1}} \mathrm{d}x \right| \leq \int_{-1/n}^{1/n} |t(x)| \underbrace{\left| \mathrm{e}^{\frac{1}{n^2x^2-1}} \right|}_{\leq 1/\mathrm{e}} \mathrm{d}x \leq \frac{1}{\mathrm{e}} \int_{-1/n}^{1/n} |t(x)| \mathrm{d}x$$

da cui si ha la contraddizione

$$1 \leq \int_{-1/n}^{1/n} |t(x)| \mathrm{d}x \xrightarrow{n \to \infty} 0 \quad \lightning$$

□

5.1.1 Funzionali lineari e continui

È necessario adesso introdurre una classe importante di oggetti, detti **funzionali**.

Definizione 5.6. *Un* funzionale $F : V_{\mathbb{K}} \to \mathbb{K}$ *è un'applicazione che mappa uno spazio vettoriale V su $\mathbb{K}$ nel campo $\mathbb{K}$ ad esso associato.*

Particolarmente interessanti per la trattazione che verrà sviluppata nel seguito sono i funzionali lineari.

Definizione 5.7. *Un funzionale F è detto* **lineare** *se $F[\alpha f + \beta g] = \alpha F[f] + \beta F[g]$.*

Esempio 5.2. Un esempio di funzionale lineare è l'applicazione $F : D \subseteq \mathscr{L}^2(\mathbb{R}) \to \mathbb{C}$ definita come $F_{\varphi_0}[f] = \langle \varphi_0, f \rangle$, con $\varphi_0 \in \mathscr{L}^2(\mathbb{R})$ fissato. Se invece $\varphi_0 \notin \mathscr{L}^2(\mathbb{R})$, allora il prodotto scalare è ben definito se $f \in \mathcal{D}(\mathbb{R})$. Quindi, in generale il dominio di questo funzionale è un sottoinsieme di $\mathscr{L}^2(\mathbb{R})$. Un altro esempio di funzionale è $F_{x_0}[f] = f(x_0) \ \forall f(x) \in C^0(\mathbb{R}) \cap \mathscr{L}^2(\mathbb{R})$, che ha un comportamento analogo alla *delta* di Dirac. Si ha che $\mathscr{S}(\mathbb{R}) \subset \operatorname{dom} F_{x_0} \subset \mathscr{L}^2(\mathbb{R})$.

Esempio 5.3. Un oggetto come $F_{\varphi_0}[f] = \varphi_0 + f$ non rappresenta un funzionale. Esso infatti mappa H in se stesso e inoltre non è difficile verificare che non risulti lineare.

Analogamente a quanto fatto con gli operatori, si introduce un concetto di **continuità** per un funzionale.

Definizione 5.8. *Un funzionale F si dice* continuo *in $\mathcal{D}(\mathbb{R})$ se $\forall \{\varphi_j(x)\} \subset \mathcal{D}(\mathbb{R}) : \varphi_j \xrightarrow{\tau_D} \varphi$ risulta $F[\varphi_j] \xrightarrow{\mathbb{C}} F[\varphi]$. Analogamente, un funzionale F si dice* continuo *in $\mathscr{S}(\mathbb{R})$ se $\forall \{\varphi_j(x)\} \subset \mathscr{S}(\mathbb{R})$: $\varphi_j \xrightarrow{\tau_S} \varphi$ risulta $F[\varphi_j] \xrightarrow{\mathbb{C}} F[\varphi]$.*

Si può quindi porre l'importante definizione di **distribuzione**.

Definizione 5.9. *Si chiama* **distribuzione** *ogni funzionale continuo in $\mathcal{D}(\mathbb{R})$. L'insieme delle distribuzioni è indicato con $\mathcal{D}'(\mathbb{R})$. Si chiama inoltre distribuzione* **temperata** *ogni funzionale continuo in $\mathscr{S}(\mathbb{R})$. L'insieme delle distribuzioni temperate è indicato con $\mathscr{S}'(\mathbb{R})$.*

Ricordiamo che lo spazio V^* formato da tutti i funzionali lineari costruiti a partire da uno spazio vettoriale lineare V viene detto **spazio duale** associato allo spazio vettoriale V. Pertanto, nel contesto delle distribuzioni, si può affermare che $\mathcal{D}'(\mathbb{R})$ sia lo spazio duale associato allo spazio di Dirichlet $\mathcal{D}(\mathbb{R})$ (analogamente per $\mathscr{S}'(\mathbb{R})$).

Esempio 5.4. Non è quindi difficile a questo punto dimostrare che la $\delta(x - x_0)$ è una distribuzione. Infatti esso è banalmente un funzionale lineare, come prima osservato. Sia $\{\varphi_j\} \subset \mathcal{D}(\mathbb{R})$ tale che $\varphi_j \xrightarrow{\tau_D} \varphi$. Si deve mostrare che $F_{x_0}[\varphi_j] = \varphi_j(x_0)$ tende a $F_{x_0}[\varphi] = \varphi(x_0)$ nella norma definita in $\mathbb{C}$. Poiché la convergenza nella topologia τ_D equivale, fra le varie cose, alla convergenza uniforme di

$\{\varphi_j\}$, essa implica in particolare la convergenza puntuale in x_0. Allora $F[\varphi_j] \xrightarrow{C} F[\varphi]$ e quindi la *delta* è una distribuzione. Inoltre, si può mostrare che essa è anche una distribuzione temperata. Infatti, anche la convergenza nella topologia τ_S comporta la convergenza uniforme di $\{\varphi_j\}$.

Proposizione 5.2. *Sia $\varphi_0 \in \mathscr{L}^2(\mathbb{R})$. Allora $F_{\varphi_0} = \langle \varphi_0, f \rangle$ è una distribuzione.*

Dimostrazione. Presa $f_n(x) \xrightarrow{\tau_D} f(x)$, allora si ha $\left|F_{\varphi_0}[f_n] - F_{\varphi_0}[f]\right| = |\langle \varphi_0, f_n \rangle - \langle \varphi_0, f \rangle| = |\langle \varphi_0, f_n - f \rangle| \leq \|\varphi_0\| \cdot \|f_n - f\|$. Ma risulta anche che $\|f_n - f\|^2 = \int_{\mathbb{R}} |f_n - f|^2 \, dx = \int_I |f_n - f|^2 \, dx \leq \mu(I) \sup |f_n - f|^2 \to 0$, dove I è l'intervallo compatto in cui è contenuto il supporto di ogni elemento della successione, in quanto $f_n \rightrightarrows f$ nella topologia τ_D. Quindi, $\left|F_{\varphi_0}[f_n] - F_{\varphi_0}[f]\right| \to 0$ e tale oggetto è una distribuzione. □

Ad ogni funzione di $\mathscr{L}^2(\mathbb{R})$ è possibile associare una distribuzione, che verrà detta **regolare**. In questo senso si dice, con un certo abuso di notazione, che $\mathscr{L}^2(\mathbb{R}) \subset \mathcal{D}'(\mathbb{R})$, non intendendo con questo che ogni funzione di $\mathscr{L}^2(\mathbb{R})$ sia una distribuzione, ma che si possa associargliene una. In generale vale anche che $\mathscr{S}'(\mathbb{R}) \subset \mathcal{D}'(\mathbb{R})$.

5.2 Derivabilità e altre proprietà

Le seguenti definizioni si possono porre con naturalezza.

Definizione 5.10. *Date due distribuzioni T_1, T_2, la loro* somma *è definita come $(T_1 + T_2)[f] = T_1[f] + T_2[f] \, \forall f \in \mathcal{D}(\mathbb{R})$ (o anche $\mathscr{S}(\mathbb{R})$). Due distribuzioni sono uguali se, e solo se, $T_1[f] = T_2[f] \, \forall f \in \mathcal{D}(\mathbb{R})$, ovvero hanno la stessa azione su tutte le funzioni test.*

Nel caso delle distribuzioni, è possibile definire un concetto di derivazione debole, che generalizza ed estende la derivazione ordinaria. Per fare questo, si premette un risultato generale sui funzionali lineari in spazi di Hilbert. Questa proposizione rappresenta una formulazione del **teorema di rappresentazione di Riesz**.

Teorema 5.5 (Rappr. di Riesz). *Dato uno spazio di Hilbert H dotato del prodotto scalare $\langle \cdot, \cdot \rangle$, tutti e soli i funzionali lineari di H sono della forma $L[x] = \langle c, x \rangle$, dove $c \in H$.*

Quindi, ad esempio, tutti gli elementi dello spazio duale di $\mathscr{L}^2(\mathbb{R})$ si possono scrivere nella forma $L[f] = \int_{\mathbb{R}} \overline{g(x)} f(x) dx$. Si è già mostrato che, restringendo il dominio a funzioni sufficientemente regolari, come funzioni di $\mathscr{S}(\mathbb{R})$, si può far cadere l'ipotesi che $g(x)$ sia una funzione. In questo modo, questa proposizione assicura che l'azione di ogni distribuzione avente per dominio $\mathscr{S}(\mathbb{R})$ o un suo sottoinsieme possa essere rappresentata in forma integrale come

$$F_\psi[f] = \langle \psi, f \rangle = \int_{\mathbb{R}} \psi(x) f(x) \, dx \, .$$

Partendo da questo fatto, si può adesso introdurre un concetto di derivazione per distribuzioni temperate, noto come *derivata debole*.

Definizione 5.11. *La* **derivata debole** *di una distribuzione temperata è l'oggetto*

$$F_{\psi'}[f] = -F_\psi[f'] \qquad \forall f \in \mathscr{S}(\mathbb{R}). \tag{5.11}$$

Equivalentemente, data una distribuzione temperata associata alla funzione ψ, la derivata debole della

distribuzione è il funzionale associato alla funzione derivata ψ', definito in maniera che

$$\int_{\mathbb{R}} \psi'(x) f(x)\, dx = -\int_{\mathbb{R}} \psi(x) \frac{df}{dx}\, dx \qquad \forall f \in \mathscr{S}(\mathbb{R}), \tag{5.12}$$

ovvero in maniera da trasferire (a meno di un segno) la derivata alla funzione test.

Tale definizione è motivata dal fatto che, se $F_\psi[f]$ è una distribuzione regolare, allora

$$\int_{\mathbb{R}} \psi' f(x)\, dx = \psi(x) f(x) \Big|_{-\infty}^{+\infty} - \int_{\mathbb{R}} \psi f'\, dx \tag{5.13}$$

in quanto $f(x) \in \mathscr{S}(\mathbb{R})$. Se $\psi(x)$ è una funzione derivabile in $\mathscr{L}^2(\mathbb{R})$, allora la legge che fornisce la derivata debole rappresenta una generalizzazione della formula di integrazione per parti.

Proposizione 5.3. *La derivata debole di una distribuzione temperata è ancora una distribuzione temperata.*

Dimostrazione. Dobbiamo dimostrare che $F_{\psi'}[f]$ è un funzionale lineare e continuo. Per prima cosa, la linearità discende banalmente dalla definizione sotto forma di integrale. Infatti

$$\begin{aligned} F_{\psi'}[\alpha f + \beta g] &= -\left(\alpha F_\psi \left[\frac{df}{dx}\right] + \beta F_\psi \left[\frac{dg}{dx}\right]\right) \\ &= \alpha \left(-F_\psi \left[\frac{df}{dx}\right]\right) + \beta \left(-F_\psi \left[\frac{dg}{dx}\right]\right) = \alpha F_{\psi'}[f] + \beta F_{\psi'}[g]. \end{aligned} \tag{5.14}$$

Mostriamo adesso la continuità del funzionale. Se consideriamo una successione $\{f_j\} \subset \mathscr{S}(\mathbb{R})$ tale che $f_j \xrightarrow{\tau_S} f$, questo implica la convergenza uniforme della successione delle derivate prime $\left\{\frac{df_j}{dx}\right\}$, da cui $F_\psi\left[\frac{df_j}{dx}\right] \xrightarrow{\mathbb{C}} F_\psi\left[\frac{df}{dx}\right]$. Questo corrisponde, a meno di un segno, alla convergenza di $F_\psi[f]$, da cui la tesi. □

È possibile dimostrare che se una funzione è derivabile nel senso ordinario, allora è anche derivabile debolmente e le due derivate coincidono. Per la derivata debole si estendono le usuali regole di calcolo delle derivate, come ad esempio la *regola di Leibniz*.

Proposizione 5.4 (Regola di Leibniz). *Se F_ψ e F_φ sono due distribuzioni temperate derivabili debolmente, allora $F_\psi F_\varphi$ è derivabile debolmente e vale la regola di Leibniz*

$$(F_\psi F_\varphi)' = F_{\psi'} F_\varphi + F_\psi F_{\varphi'} \tag{5.15}$$

Ricordiamo che è convenzione indicare il funzionale lineare F_ψ più semplicemente con ψ, ovvero con l'oggetto che genera tale funzionale. Nel seguito, adotteremo anche noi questa convenzione.

Proposizione 5.5. *Se $F_\varphi \in \mathcal{D}'(\mathbb{R})$ e $\psi \in \mathcal{D}(\mathbb{R})$, allora $\psi\varphi$ è derivabile debolmente.*

Dimostrazione. Si ha che $F_{\varphi\psi}[f] = \langle \psi\varphi, f \rangle = \int_{\mathbb{R}} \psi(x)\varphi(x) f(x)\, dx = \langle \varphi, \psi f \rangle = F_\varphi[\psi f] \in \mathcal{D}'(\mathbb{R})$, visto che $f\psi \in \mathcal{D}(\mathbb{R})$. □

Una proposizione analoga vale se $F_\varphi \in \mathscr{S}'(\mathbb{R})$ e $\psi \in \mathscr{S}(\mathbb{R})$.

5.2.1 Esempi di derivate deboli

Esempio 5.5. Si calcoli la derivata debole della funzione $u(x)$, il cui grafico è riportato in Figura 5.1. La derivata ordinaria esiste dovunque meno che in $x = 0$. La derivata debole invece

$$\int_{\mathbb{R}} u'(x)f(x)\,\mathrm{d}x = -\int_{\mathbb{R}} u(x)\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x = -\int_0^{+\infty}\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x = -\,f(x)\Big|_0^{+\infty} = f(0) = \int_{\mathbb{R}} \delta(x)f(x)\,\mathrm{d}x$$

esiste ed è uguale alla *delta* di Dirac.
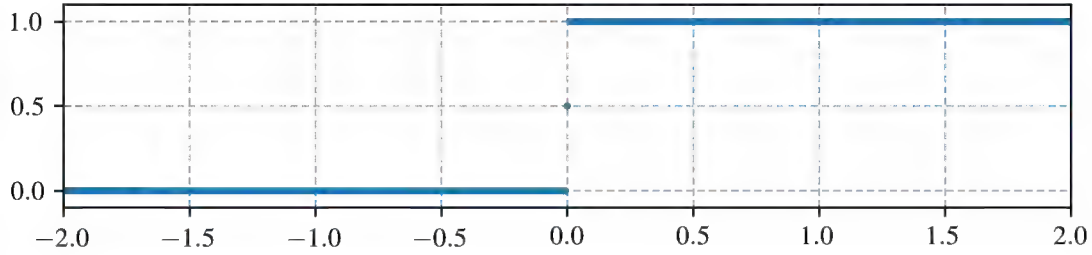
Figura 5.1: Funzione *gradino*, ovvero $u(x)$.

Esempio 5.6. Si calcoli la derivata debole della funzione[a] $\mathrm{rmp}(x) = u(x)x$, che ha il grafico riportato in Figura 5.2. La derivata ordinaria esiste ovunque meno che in $x = 0$, a causa della presenza di un punto angoloso. Per la derivata debole invece si ha

$$\int_{\mathbb{R}} \mathrm{rmp}'(x)f(x)\,\mathrm{d}x = -\int_{\mathbb{R}} \mathrm{rmp}(x)\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x = -\int_0^{+\infty} x\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x = -\left\{\cancel{xf(x)|_{\mathbb{R}}} - \int_0^{+\infty} f(x)\,\mathrm{d}x\right\}$$
$$= \int_{\mathbb{R}} u(x)f(x)\,\mathrm{d}x\,,$$

che esiste ed è uguale a $u(x)$. Notiamo che questo risultato si poteva più semplicemente ottenere tramite la regola di Leibniz, ovvero

$$\mathrm{rmp}'(x) = [u(x)x]' = u'(x)x + u(x) = \delta(x)x + u(x) = u(x), \tag{5.16}$$

ricordando che in generale $\delta(x)g(x) = g(0)$.
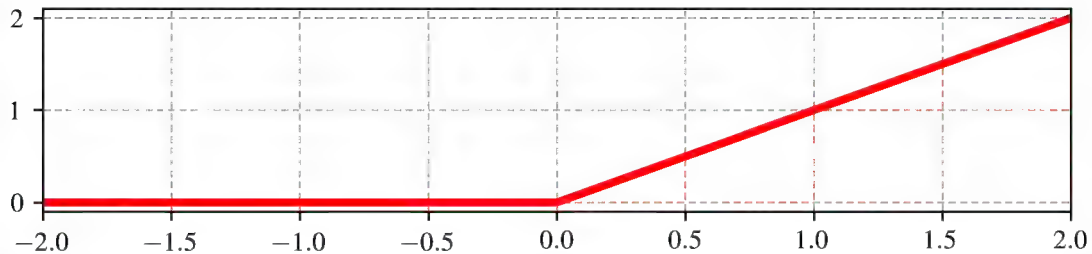
Figura 5.2: Funzione *rampa*, ovvero $\mathrm{rmp}(x)$.

[a] $u(x)$ è la funzione gradino di Heaviside.

Esempio 5.7. Si calcoli la derivata debole della funzione $\mathrm{rect}(x)$, che ha il grafico riportato in Figura 5.3. La derivata ordinaria esiste ovunque meno che nelle discontinuità in $x = \pm 1/2$. La derivata debole invece

$$\begin{aligned}\int\limits_{\mathbb{R}} (\mathrm{rect}(x))' f(x)\,\mathrm{d}x &= -\int\limits_{\mathbb{R}} \mathrm{rect}(x)\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x = -\int\limits_{-1/2}^{1/2} \frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x = -f\left(\frac{1}{2}\right) + f\left(-\frac{1}{2}\right) \\ &= \int\limits_{\mathbb{R}} \left[\delta\left(x-\frac{1}{2}\right) - \delta\left(x+\frac{1}{2}\right)\right] f(x)\,\mathrm{d}x\end{aligned}$$

esiste ed è uguale a $\delta(x - 1/2) - \delta(x + 1/2)$. In generale, la derivata debole di una funzione costante a tratti è una somma di funzioni δ, ognuna moltiplicata per un fattore pari al salto $f_+ - f_-$, dove f_+ (f_-) rappresenta il limite destro (sinistro) della funzione nel punto di discontinuità. (Questo fatto è alquanto simpatico da dimostrare, il lettore più avvezzo è invitato ad abbozzare una dimostrazione).
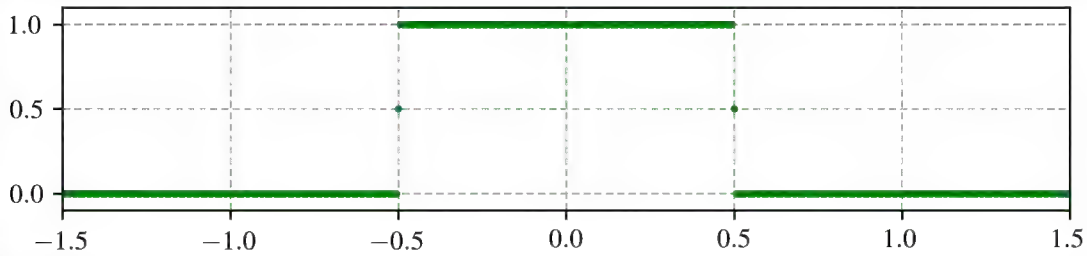
Figura 5.3: Funzione *rettangolo*, ovvero $\mathrm{rect}(x)$.

5.3 Distribuzioni come limite debole

È possibile introdurre il concetto di distribuzione anche proseguendo per una strada alternativa. Ad esempio, si definisca la successione di funzioni $\delta_n = n\,\mathrm{rect}(nx)$. Ogni funzione di tale successione è tale che $\int_{\mathbb{R}} \delta_n(x)\,\mathrm{d}x = 1$, cosa che garantisce che $\delta_n \neq 0$ quasi ovunque. Tale successione converge puntualmente a zero dovunque tranne che in $x = 0$. Tuttavia,

$$\lim_{n,\infty} \int\limits_{\mathbb{R}} \delta_n(x) f(x)\mathrm{d}x = \lim_{n,\infty} n \int\limits_{-1/2n}^{1/2n} f(x)\mathrm{d}x = \lim_{n,+\infty} f(\bar{x}_n) = f(0) = \int\limits_{\mathbb{R}} \delta(x) f(x)\mathrm{d}x \tag{5.17}$$

avendo sfruttato il teorema della media integrale, ovvero che

$$\exists\, \bar{x}_n \in \left(-\frac{1}{2n}, \frac{1}{2n}\right) \quad \text{t.c.} \quad \int\limits_{-1/2n}^{1/2n} f(x)\,\mathrm{d}x = \frac{f(\bar{x}_n)}{n}. \tag{5.18}$$

Quindi, la successione converge in qualche modo alla *delta* di Dirac, nel senso che sotto il segno di integrale $\delta(x)$ e il limite della successione delle δ_n hanno la stessa azione su tutte le funzioni test. Questo motiva la seguente definizione.

Definizione 5.12. *Si dice che una successione di funzioni* $\{\varphi_n\} \subset \mathscr{S}(\mathbb{R})$ converge debolmente *a una distribuzione temperata* φ *se*

$$\lim_{n,+\infty} \int_{\mathbb{R}} \varphi_n(x) f(x)\, \mathrm{d}x = \int_{\mathbb{R}} \varphi(x) f(x)\, \mathrm{d}x \quad \forall f \in \mathscr{S}(\mathbb{R}) \tag{5.19}$$

Vale il seguente teorema che garantisce che il procedimento precedentemente illustrato per ottenere la *delta* di Dirac è del tutto generale.

Teorema 5.6. *Tutte le distribuzioni temperate possono essere ottenute come limite debole di successioni di funzioni di* $\mathscr{S}(\mathbb{R})$.

La scelta della successione approssimante non è unica. Ad esempio, nel caso della *delta* di Dirac vale la seguente proposizione.

Proposizione 5.6. *Data una qualsiasi* $\chi(x)$ *pari e non negativa, tale che* $\int_{\mathbb{R}} \chi(x)\, \mathrm{d}x = 1$, *la successione* $\delta_n(x) = n\chi(nx)$ *converge debolmente alla* delta *di Dirac.*

Esempio 5.8. Un altro modo di costruire una successione approssimante per la *delta* di Dirac è sfruttare il Teorema 1.2 per ottenere una rappresentazione in serie[a]. Sia $\{\mathrm{e}_n(x)\}$ una base ortonormale e totale di $\mathscr{L}^2(\mathbb{R})$. Tramite una serie di passaggi formali, si ottiene che

$$\begin{aligned}
\langle f, g\rangle &= \sum_n \langle f, \mathrm{e}_n\rangle \langle \mathrm{e}_n, g\rangle = \sum_n \int_{\mathbb{R}} \overline{f(x)}\, \mathrm{e}_n(x)\, \mathrm{d}x \sum_n \int_{\mathbb{R}} \overline{\mathrm{e}_n(y)}\, g(y)\, \mathrm{d}x \\
&= \int_{\mathbb{R}^2} \mathrm{d}x\, \mathrm{d}y\, \overline{f(x)} g(y) \underbrace{\sum_n \mathrm{e}_n(x)\overline{\mathrm{e}_n(y)}}_{\Delta(x,y)} = \int_{\mathbb{R}^2} \mathrm{d}x\, \mathrm{d}y\, \overline{f(x)} g(y) \Delta(x,y) \\
&= \int_{\mathbb{R}} \overline{f(x)} \left(\int_{\mathbb{R}} \Delta(x,y) g(y)\, \mathrm{d}y \right) \mathrm{d}x
\end{aligned} \tag{5.20}$$

Affinché valga l'uguaglianza deve valere che $g(x) = \int_{\mathbb{R}} \Delta(x,y) g(y)\, \mathrm{d}x$ e quindi $\Delta(x,y) = \delta(x - y)$. Allora, la *delta* di Dirac può essere ottenuta come somma della serie $\Delta(x,y) = \sum_n \mathrm{e}_n(x)\overline{\mathrm{e}_n(y)}$.

[a] Infatti, ogni serie è in realtà una successione, più nello specifico la successione delle somme parziali.

5.4 Proprietà della delta di Dirac

Qui di seguito sono elencate, e in alcuni casi dimostrate, varie proprietà della delta di Dirac che risultano estremamente utili nelle applicazioni.

Proposizione 5.7. $\delta(-x) = \delta(x)$, *ovvero la* delta *è una distribuzione pari.*

Dimostrazione. $\int_{\mathbb{R}} \delta(-x) f(x)\, \mathrm{d}x = -\int_{+\infty}^{-\infty} \delta(y) f(-y)\, \mathrm{d}x = \int_{\mathbb{R}} \delta(y) f(-y)\, \mathrm{d}x = f(0) = \int_{\mathbb{R}} \delta(x) f(x)\, \mathrm{d}x \;\therefore\; \delta(-x) = \delta(x)$. □

D'altronde, la *delta* è ottenibile come limite debole di successioni di funzioni pari. È quindi naturale che anch'essa debba essere pari. Così come la derivata di una funzione pari è una funzione dispari, anche per la *delta* vale che la sua derivata debole ha parità definita.

Proposizione 5.8. *La derivata debole della $\delta(x)$ è una distribuzione dispari, ovvero tale che $\delta'(-x) = -\delta'(x)$.*

Dimostrazione. Si presenta qui una giustificazione prettamente formale di questo risultato. Si ha infatti che

$$\delta'(x) = \frac{i}{2\pi} \int_{\mathbb{R}} p e^{ipx} \, dp \tag{5.21}$$

da cui

$$\delta'(-x) = \frac{i}{2\pi} \int_{\mathbb{R}} p e^{-ipx} \, dp = \frac{i}{2\pi} \int_{+\infty}^{-\infty} y e^{iyx} \, dx = -\frac{i}{2\pi} \int_{\mathbb{R}} y e^{iyx} \, dx = -\delta'(x) \tag{5.22}$$

da cui la tesi. □

Proposizione 5.9. *Sia $a \neq 0$. Allora $\delta(a\,x) = \frac{1}{|a|}\delta(x)$.*

Dimostrazione. Effettuando la posizione $ax = y$:

$$\begin{aligned}
\int_{\mathbb{R}} \delta(ax) f(x) \, dx &= \begin{cases} \displaystyle\int_{\mathbb{R}} \frac{1}{a}\delta(y) f\left(\frac{y}{a}\right) dy & a > 0 \\[2ex] \displaystyle\int_{+\infty}^{-\infty} \frac{1}{a}\delta(y) f\left(\frac{y}{a}\right) dy & a < 0 \end{cases} \\
&= \begin{cases} \displaystyle\frac{1}{a}\int_{\mathbb{R}} \delta(y) f\left(\frac{y}{a}\right) dy & a > 0 \\[2ex] \displaystyle\frac{1}{-a}\int_{\mathbb{R}} \delta(y) f\left(\frac{y}{a}\right) dy & a < 0 \end{cases} \\
&= \frac{1}{|a|}\int_{\mathbb{R}} \delta(y) f\left(\frac{y}{a}\right) dy = \frac{1}{|a|} f(0) = \frac{1}{|a|}\int_{\mathbb{R}} \delta(x) f(x) \, dx
\end{aligned} \tag{5.23}$$

da cui la tesi. □

Proposizione 5.10. *Data una qualsiasi $g(x) \in \mathscr{S}(\mathbb{R})$, si ha che $\delta(x - x_0) g(x) = g(x_0)\delta(x - x_0)$.*

Dimostrazione. $\int_{\mathbb{R}} \delta(x - x_0) g(x) f(x) \, dx = \int_{\mathbb{R}} \delta(x - x_0) \left(g(x) f(x)\right) dx = g(x_0) f(x_0) = g(x_0) \int_{\mathbb{R}} \delta(x - x_0) f(x) \, dx = \int_{\mathbb{R}} \left[g(x_0)\delta(x - x_0)\right] f(x) \, dx$, da cui segue l'asserto. □

La Proposizione 5.9 è un caso particolare di una formula più generale.

Proposizione 5.11. *Data una funzione $g(x)$, indicati con $\{x_j\}_{j=1}^{n}$ gli zeri di tale funzione, ovvero i punti tali che $g(x_j) = 0$, si ha che*

$$\delta(g(x)) = \sum_{j=1}^{n} \frac{\delta(x - x_j)}{|g'(x_j)|} \tag{5.24}$$

Proposizione 5.12. *Se quando si considera l'azione di $\delta(x-a)$ sotto il segno di integrale, l'integrazione è svolta su un intervallo finito, allora*

$$\int_a^b \delta(x-x_0)f(x)\,\mathrm{d}x = \begin{cases} f(x_0) & x_0 \in]a,b[\\ 0 & x_0 \notin [a,b] \\ \frac{f(x_0)}{2} & x_0 = a \vee x_0 = b \end{cases} \tag{5.25}$$

Tutte le proprietà della $\delta(x)$ possono essere ricavate, più o meno semplicemente, utilizzando l'espressione formale

$$\delta(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ipx}\mathrm{d}p$$

già definita in Equazione 5.2.

5.5 Cenni sulla parte principale di Cauchy

Introduciamo adesso una distribuzione non regolare, ottenuta per dare significato agli integrali che divergono a causa della presenza di un polo semplice lungo il cammino di integrazione. Data una $f(x) \in \mathcal{D}(\mathbb{R})$ con $f(x_0) \neq 0$, sia I l'integrale definito come

$$I = \int_{-\infty}^{+\infty} \frac{f(x)}{x-x_0}\mathrm{d}x, \tag{5.26}$$

che in generale non esiste a causa della divergenza dell'integrando in $x = x_0$. In ogni caso, possiamo assegnare un qualche valore a tale espressione definendo il cosiddetto **valore principale**, o **parte principale** di Cauchy, dell'integrale I.

Definizione 5.13. *Dato $\epsilon > 0$, si definisce valore principale di Cauchy il risultato del limite*

$$\mathcal{P}\left(\frac{1}{x-x_0}\right)[f] = \text{P.V.}\int_{-\infty}^{+\infty} \frac{f(x)}{x-x_0}\mathrm{d}x := \lim_{\epsilon,0^+}\left[\int_{-\infty}^{x_0-\epsilon} \frac{f(x)}{x-x_0}\mathrm{d}x + \int_{x_0+\epsilon}^{+\infty} \frac{f(x)}{x-x_0}\mathrm{d}x\right]. \tag{5.27}$$

In particolare si mostra che se I esiste, esso coincide con il suo valore principale e che $\mathcal{P}\left(\frac{1}{x-x_0}\right) \in \mathcal{D}'(\mathbb{R})$. Vale la seguente formula.

Teorema 5.7 (di Sochocki-Plemelj). *Il valore principale di Cauchy si può rappresentare come*

$$\mathcal{P}\left(\frac{1}{x-x_0}\right)[f] = \lim_{\epsilon,0^+} \frac{1}{x-x_0 \pm i\epsilon} \pm i\pi\delta(x-x_0) \tag{5.28}$$

dove il segno dipende dal cammino di integrazione seguito, se deformato nel semipiano delle parti immaginarie positive o negative.

5.6 Funzioni di Green

Le distribuzioni rivestono un ruolo importantissimo nella teoria delle equazioni differenziali. In molti casi, soprattutto quando si trattano problemi differenziali alle derivate parziali, non si riesce a trovare una soluzione, non foss'altro perché le condizioni di regolarità che tale soluzione

dovrebbe avere non riescono ad essere soddisfatte da alcuna funzione.
Si può approcciare il problema nella sua cosiddetta *formulazione debole*. In generale, un'equazione differenziale può essere pensata come un'espressione del tipo

$$\mathcal{L}\left[\Phi(x,t)\right] = 0, \tag{5.29}$$

dove $\mathcal{L}$ è un operatore differenziale. Allora, facendo il prodotto scalare di tale espressione con una funzione di $\mathscr{S}(\mathbb{R})$, si ha che

$$0 = \langle \mathcal{L}\Phi, f\rangle = \left\langle \Phi, \mathcal{L}^{\dagger} f\right\rangle \iff \mathcal{L}^{\dagger} f = 0. \tag{5.30}$$

Risolta l'equazione con $\mathcal{L}^{\dagger}$ in $\mathscr{S}(\mathbb{R})$, si può quindi determinare $\Phi(x,t)$, che in generale non deve essere necessariamente una funzione, ma può essere una distribuzione. Tale oggetto viene detta **soluzione debole** del problema.

Senza scendere in ulteriori dettagli in queste questioni, si affronta qui più semplicemente il problema della risoluzione di un'equazione differenziale ordinaria lineare del secondo ordine col metodo della **funzione di Green**. In generale, una ODE lineare di ordine 2 si scrive come

$$y''(x) + p(x)y'(x) + q(x)y(x) = f(x) \tag{5.31}$$

che, introducendo l'operatore differenziale lineare $\mathcal{L} = \frac{\mathrm{d}^2}{\mathrm{d}x^2} + p(x)\frac{\mathrm{d}}{\mathrm{d}x} + q(x)$, diventa $\mathcal{L}[y] = f(x)$. È noto che in questo caso l'integrale generale si ottiene sommando la soluzione dell'omogenea associata con una soluzione particolare. Il metodo della funzione di Green permette di determinare la soluzione particolare.

Definizione 5.14. *Si indica con* funzione di Green *la funzione* $G(\mathbf{x})$ *tale che*

$$\mathcal{L}[G(\mathbf{x})] = \delta(\mathbf{x}).$$

Se essa esiste, allora non è difficile verificare che $y_{PT}(x) = \int_{\mathbb{R}} G(x-s)\, f(s)\, \mathrm{d}s$. Infatti

$$\mathcal{L}[y_{PT}] = \mathcal{L}\left[\int_{\mathbb{R}} G(x-s)\, f(s)\, \mathrm{d}s\right] = \int_{\mathbb{R}} \mathcal{L}[G(x-s)] f(s)\, \mathrm{d}s = \int_{\mathbb{R}} \delta(x-s) f(s)\, \mathrm{d}s = f(x)$$

Si danno ora due esempi di applicazione di questa tecnica.

Esempio 5.9. (ODE) Sia data l'equazione differenziale $y'' - 2y = x$. Non è difficile verificare che la soluzione particolare è nella forma $y_P(x) = -x/2$, sfruttando il principio di identità dei polinomi. Per applicare la tecnica di *Green*, si scrive la funzione $G(x)$ in termini della sua trasformata di Fourier.

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{G}(p)\, e^{ipx}\, \mathrm{d}p\,.$$

Applicando ad essa l'operatore differenziale lineare $L = \frac{\mathrm{d}^2}{\mathrm{d}x^2} - 2$, associato all'equazione in studio,

si ottiene

$$L[G(x)] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{G}(p) L[e^{ipx}] \, dp = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{G}(p)(-p^2 - 2) \, e^{ipx} \, dp$$
$$= \delta(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ipx} \, dp,$$

da cui

$$\hat{G}(p) = -\frac{1}{\sqrt{2\pi}} \frac{1}{2+p^2}.$$

Poiché tale funzione è in $\mathscr{L}^2(\mathbb{R})$, è possibile antitrasformare senza indugio, ottenendo

$$G(x) = -\frac{1}{2\sqrt{2}} \, e^{-\sqrt{2}|x|}.$$

Allora

$$y_P(x) = \int_{\mathbb{R}} G(x-s) \, f(s) \, ds = -\frac{1}{2\sqrt{2}} \int_{\mathbb{R}} e^{-\sqrt{2}|x-s|} \, s \, ds = -\frac{x}{2}.$$

5.6.1 Scattering quantistico

La tecnica di *Green* permette di dare una rappresentazione integrale dell'equazione di Schrödinger associata al problema dello scattering in meccanica quantistica. Si ha a che fare con un fascio di particelle, moventesi lungo la direzione z, che incidono su un centro di scattering, ad esempio un atomo, uno ione etc. La forma della funzione d'onda dopo l'interazione con il centro di scattering si ricava dall'equazione di *Schrödinger*

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}),$$

che può essere riscritta nella forma

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = U(\mathbf{r})\psi(\mathbf{r}), \tag{5.32}$$

con $k^2 = \frac{2mE}{\hbar^2}$ e $U(\mathbf{r}) = \frac{2m}{\hbar^2} V(\mathbf{r})$.

Si può determinare una soluzione particolare di (5.32) tramite il metodo di *Green*. In questo caso, $G(\mathbf{r})$ è la funzione tale che

$$(\nabla^2 + k^2) \, G(\mathbf{r}) = \delta(\mathbf{r})$$

dove $\delta(\mathbf{r}) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{i\mathbf{p}\cdot\mathbf{r}} d^3p$ è la versione tridimensionale della *delta* di Dirac. Detta poi $\Phi(\mathbf{r})$ la soluzione dell'omogenea associata a (5.32), si ha che l'integrale generale è nella forma

$$\boxed{\psi(\mathbf{r}) = \Phi(\mathbf{r}) + \int_{\mathbb{R}^3} G(\mathbf{r} - \mathbf{r}_1) \, U(\mathbf{r}_1) \, \psi(\mathbf{r}_1) \, d^3 r_1} \tag{5.33}$$

che è l'equazione integrale associata all'equazione differenziale di prima, ovvero una sua riscrittura. Data la formulazione del problema, non è difficile convincersi che $\Phi(\mathbf{r}) = e^{i\mathbf{k}_0 \cdot \mathbf{r}} = e^{ikz}$. Una scrittura del tipo (5.33) è difficilmente esplicitabile quando si considerano potenziali *realistici*, ovvero ad esempio anarmonici, non aventi simmetria centrale etc. La cosa allora che è possibile fare

per ottenere comunque una soluzione approssimata del problema è iterare l'integrazione secondo la formula

$$\psi_n(\mathbf{r}) = \Phi(\mathbf{r}) + \int_{\mathbb{R}^3} G(\mathbf{r}-\mathbf{r_1})\, U(\mathbf{r_1})\, \psi_{n-1}(\mathbf{r_1})\, \mathrm{d}^3 r_1 \tag{5.34}$$

che viene applicata iterativamente finché $\|\psi_n - \psi_{n-1}\| < \epsilon$, dove ϵ è un errore arbitrariamente fissato. In questo approccio **pertubativo**, l'approssimazione al primo ordine viene detta *di Born*. Tale procedura in generale non converge e pertanto la sua applicabilità va verificata caso per caso. Si procede ora a determinare la forma di $G(\mathbf{r})$. Si ha che

$$\begin{aligned}(\nabla^2 + k^2)G(\mathbf{r}) &= \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} (\nabla^2_{\mathbf{r}} + k^2) \mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}} \hat{G}(\mathbf{p})\, \mathrm{d}\mathbf{p} = \\ &= \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} (-|\mathbf{p}|^2 + k^2) \mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}} \hat{G}(\mathbf{p})\, \mathrm{d}\mathbf{p} = \delta(\mathbf{r})\end{aligned}$$

da cui $\hat{G}(\mathbf{p}) = \dfrac{1}{(2\pi)^{3/2}} \dfrac{1}{k^2 - p^2}$. Da questo si ricava, passando in coordinate sferiche:

$$G(\mathbf{r}) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{\mathrm{e}^{-i\mathbf{p}\cdot\mathbf{r}}}{k^2 - p^2} \mathrm{d}^3 p = \frac{1}{(2\pi)^3} \int_0^{2\pi} \mathrm{d}\phi \int_0^{+\infty} \mathrm{d}p \frac{p^2}{k^2 - p^2} \int_0^{\pi} \sin(\theta) \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{r}} \mathrm{d}\theta$$

$$= \frac{1}{(2\pi)^2} \int_0^{+\infty} \mathrm{d}p \frac{p^2}{k^2 - p^2} \int_0^{\pi} \sin\theta \mathrm{e}^{-ipr\cos\theta} \mathrm{d}\theta = -\frac{1}{(2\pi)^2} \int_0^{+\infty} \mathrm{d}p \frac{p^2}{k^2 - p^2} \int_{-1}^{1} \mathrm{d}x \mathrm{e}^{-iprx}$$

$$= \frac{1}{(2\pi)^2} \int_0^{+\infty} \mathrm{d}p \frac{p^2}{k^2 - p^2} \left[\frac{\mathrm{e}^{-iprx}}{ipr} \right]_1^{-1} = \frac{1}{(2\pi)^2} \int_0^{+\infty} \mathrm{d}p \frac{p^2}{k^2 - p^2} \frac{\mathrm{e}^{ipr} - \mathrm{e}^{-ipr}}{ipr}.$$

Per applicare il teorema dei residui, cerchiamo di estendere l'integrale a $\mathbb{R}$:

$$\frac{1}{4i\pi^2 r} \int_0^{+\infty} \mathrm{d}p \frac{p\,(\mathrm{e}^{ipr} - \mathrm{e}^{-ipr})}{k^2 - p^2} \underbrace{=}_{-p=:q \text{ nel secondo}} \frac{1}{4i\pi^2 r} \left[\int_0^{+\infty} \mathrm{d}p \frac{p\mathrm{e}^{ipr}}{k^2 - p^2} + \int_{-\infty}^{0} \mathrm{d}q \frac{q\mathrm{e}^{irq}}{k^2 - q^2} \right] \tag{5.35}$$

che è uguale, essendo la variabile di integrazione muta, a

$$G(\mathbf{r}) = \frac{i}{4\pi^2 r} \int_{-\infty}^{+\infty} \frac{p\,\mathrm{e}^{ipr}}{p^2 - k^2} \mathrm{d}p\,. \tag{5.36}$$

Il calcolo di questo integrale è abbastanza problematico in quanto non esiste, a causa della presenza di due singolarità lungo il cammino di integrazione, come visto in precedenza nel caso della parte principale di Cauchy. Si possono però adottare diverse strategie formali che permettono, in un certo senso, di tirar fuori dei risultati fisicamente sensati. Si introducano $k_\pm := k \pm i\epsilon$. Allora, poiché r è definito positivo, si ha che

$$G_+(\mathbf{r}) = \frac{i}{4\pi^2 r} \int_{-\infty}^{+\infty} \frac{p\,\mathrm{e}^{ipr}}{p^2 - k_+^2} \mathrm{d}p = \frac{i}{4\cancel{\pi^2} r} \cancel{2\pi}\, i \,\mathrm{Res}\left\{ \frac{p\,\mathrm{e}^{ipr}}{p^2 - k_+^2}, k_+ \right\} = -\frac{\mathrm{e}^{ik_+ r}}{4\pi r}.$$

Analogamente, si ottiene

$$G_{-}(\mathbf{r}) = -\frac{\mathrm{e}^{-ik\ r}}{4\pi r}, \qquad r \equiv |\mathbf{r}|.$$

Le funzioni di Green associate a queste due espressioni si ottengono passando al limite per $\epsilon \to 0$, ottenendo

$$G_{out}(\mathbf{r}) = -\frac{\mathrm{e}^{ikr}}{4\pi r} \quad \mathrm{e} \quad G_{in}(\mathbf{r}) = -\frac{\mathrm{e}^{-ikr}}{4\pi r}, \tag{5.37}$$

che sono dette rispettivamente *outgoing* e *ingoing*. Il fatto che le due diverse deformazioni producano soluzioni diverse deriva proprio dal fatto che l'integrale di partenza in realtà non esiste. Il significato della nomenclatura di queste soluzioni può essere spiegato analizzando il comportamento asintotico della soluzione (5.33), ovvero per $r \to +\infty$. Nell'ipotesi naturale che il potenziale di interazione abbia raggio d'azione finito, si ha

$$\psi_{out}(\mathbf{r}) \simeq \Phi(\mathbf{r}) - \frac{\mathrm{e}^{ikr}}{4\pi r} \int_{\mathbb{R}^3} U(\mathbf{r_1})\psi_{out}(\mathbf{r_1})\,\mathrm{d}^3 r_1 \tag{5.38}$$

$$\psi_{in}(\mathbf{r}) \simeq \Phi(\mathbf{r}) - \frac{\mathrm{e}^{-ikr}}{4\pi r} \int_{\mathbb{R}^3} U(\mathbf{r_1})\psi_{in}(\mathbf{r_1})\,\mathrm{d}^3 r_1 \tag{5.39}$$

Quindi, la soluzione generale asintoticamente si comporta come la somma di un'onda piana e di un'onda sferica, uscente o entrante a seconda del segno dell'esponente.

Capitolo 6

Elementi di teoria dei gruppi

I gruppi sono utilizzati in tutte le branche della matematica e in molti problemi della fisica e delle altre scienze; spesso servono a catturare la simmetria intrinseca di altre strutture. Nel seguito, introdurremo il concetto di gruppo e daremo alcuni elementi riguardo ai cosiddetti gruppi continui.

6.1 Definizioni e primi esempi

Diamo la definizione di gruppo.

Definizione 6.1. *Dato un campo* $\mathbb{K}$*, un* **gruppo** *è un insieme* $\mathcal{G} = \{g_\alpha \,|\, \alpha \in \mathbb{K}\}$ *dotato di una legge di composizione dei suoi elementi tale che*

1. *l'insieme è* **chiuso** *rispetto a tale legge, ovvero* $\forall\, \alpha, \beta \in \mathbb{K}\, \exists\, \gamma \in \mathbb{K} \,|\, g_\alpha g_\beta = g_\gamma$*;*
2. *la legge è* **associativa***:* $g_\alpha(g_\beta g_\gamma) = (g_\alpha g_\beta)g_\gamma \;\forall\, \alpha, \beta, \gamma \in \mathbb{K}$*;*
3. *esiste l'***elemento neutro***, ovvero* $\exists\, g_0 \in \mathcal{G} \,|\, g_\alpha g_0 = g_0 g_\alpha = g_\alpha \;\forall\, \alpha \in \mathbb{K}$*;*
4. *esiste l'***inverso** *di ogni elemento, ovvero* $\forall\, g_\alpha \in \mathcal{G}\; \exists\, g_\beta \in \mathcal{G} \,|\, g_\alpha g_\beta = g_\beta g_\alpha = g_0$*.*

Esempio 6.1. L'insieme $B(H)$ che contiene gli operatori limitati definiti sullo spazio di Hilbert H non è un gruppo rispetto all'operazione di moltiplicazione degli operatori in quanto non tutti gli operatori limitati ammettono un operatore inverso. Restringendosi però al sottospazio $\mathcal{B} = \{\beta \in B(H) : \exists \beta^{-1} \in H\}$, esso risulta essere un gruppo rispetto a tale operazione.

Introduciamo adesso un'importante classe di gruppi.

Definizione 6.2. *Dato un gruppo, se la legge di composizione soddisfa la proprietà di* commutatività*, ovvero se* $\forall\, \alpha, \beta \in \mathbb{K}$ *si ha* $g_\alpha g_\beta = g_\beta g_\alpha$*, allora il gruppo viene detto* **abeliano** *o* **commutativo***.*

Definizione 6.3. *L'***ordine** *di un gruppo è pari al numero di elementi contenuti nel gruppo.*

È immediato mostrare che in un gruppo l'elemento neutro e l'inverso sono unici.

Proposizione 6.1. *Dato un gruppo* $\mathcal{G}$*, allora l'elemento neutro* g_0 *è unico e, inoltre,* $\forall\, g_\alpha \in \mathcal{G}$ *è unico il suo inverso* g_α^{-1}*.*

Dimostrazione. Si supponga l'esistenza di due elementi neutri g_0 e g_0' tali che $g_\alpha g_0 = g_0 g_\alpha = g_\alpha$ e $g_\alpha g_0' = g_0' g_\alpha = g_\alpha$. Si scelga ad esempio la seconda di queste relazioni. Poiché questa vale $\forall\,\alpha \in \mathbb{K}$, allora certamente vale anche per $\alpha = 0$, da cui $g_0 = g_0 g_0' = g_0'$, dove la seconda uguaglianza è motivata dal fatto che g_0 è elemento neutro. Quindi, tale elemento è unico. Analogamente si procede per mostrare l'unicità dell'inverso. □

Proposizione 6.2. *L'inverso di un prodotto di elementi del gruppo è* $(g_\alpha g_\beta)^{-1} = g_\beta^{-1} g_\alpha^{-1}$. *Inoltre,* $(g_\alpha^{-1})^{-1} = g_\alpha$.

Vengono di seguito elencati alcuni esempi notevoli di gruppi.

Esempio 6.2. $\mathcal{G}_1 = \mathbb{Q} \setminus \{0\}$ è un gruppo abeliano rispetto all'usuale operazione di moltiplicazione.

Esempio 6.3. $\mathcal{G}_2 = \mathbb{R}$ è un gruppo abeliano rispetto alla somma abituale. Per questo, è detto essere un *gruppo additivo*.

Esempio 6.4. $\mathcal{G}_p = \{2n, n \in \mathbb{Z}\}$, insieme formato da tutti gli interi pari, è un gruppo additivo. Al contrario, l'insieme $\mathcal{G}_d = \{2n+1\,, n \in \mathbb{Z}\}$ non è un gruppo. Infatti, non è un gruppo additivo in quanto la somma di due numeri dispari è un numero pari e pertanto non appartiene a $\mathcal{G}_d$ e, inoltre, manca l'elemento neutro rispetto alla somma.

Esempio 6.5. Un *gruppo moltiplicativo* è un insieme che risulta avere la struttura di gruppo abeliano rispetto all'usuale operazione di moltiplicazione. Un esempio interessante di un oggetto del genere è $U(1)$, ovvero l'insieme dei numeri complessi con modulo unitario

$$U(1) = \{\mathrm{e}^{i\alpha} \,|\, \alpha \in \mathbb{R}\}. \tag{6.1}$$

Infatti,

1. $\mathrm{e}^{i\alpha}\mathrm{e}^{i\beta} = \mathrm{e}^{i(\alpha+\beta)} \in U(1)$;
2. valgono banalmente l'associatività e la commutatività;
3. $\mathrm{e}^{i\cdot 0} = 1$ è l'elemento neutro;
4. $\mathrm{e}^{-i\alpha}$ è l'inverso di $\mathrm{e}^{i\alpha}$.

Questo gruppo è fisicamente interessante in quanto la moltiplicazione di una funzione d'onda $\psi(\mathbf{r})$ in meccanica quantistica per un numero complesso di modulo unitario non altera l'informazione che essa contiene. In altre parole, in $U(1)$ sono contenute tutti i fattori di fase *globale* per cui è possibile moltiplicare una funzione d'onda senza modificarne il significato fisico.

Esempio 6.6. Si consideri l'insieme $\mathcal{G}_3 = \{A \in \mathbb{R}^{2\times 2} \,|\, \det A \neq 0\}$. Non è difficile verificare che esso è un gruppo non abeliano. Si noti che la condizione $\det A \neq 0$ è necessaria per garantire l'esistenza della matrice inversa, e quindi la struttura di gruppo.

Esempio 6.7. Sia T_v l'operatore, detto di traslazione, che agisce su un vettore $\mathbf{w} \in \mathbb{R}^3$ come $T_v\,\mathbf{w} = \mathbf{v} + \mathbf{w}$. L'insieme $\mathcal{G}_4 = \{T_v \,|\, \mathbf{v} \in \mathbb{R}^3\}$ ha una struttura di gruppo rispetto all'ordinaria operazione di moltiplicazione di operatori, intesa come l'applicazione successiva degli operatori. Infatti, ad

esempio, $T_{v_1}T_{v_2}\mathbf{w} = T_{v_1}(\mathbf{v_2} + \mathbf{w}) = (\mathbf{v_1} + \mathbf{v_2}) + \mathbf{w} = T_{v_1+v_2}\mathbf{v}$ e quindi $T_{v_1}T_{v_2} \in \mathcal{G}_4$.

Esempio 6.8. SO (3) è l'insieme delle rotazioni in $\mathbb{R}^3$. Esso forma un gruppo non abeliano.

Si possono combinare più gruppi per ottenere gruppi di ordine maggiore.

Definizione 6.4. *Il* **prodotto diretto** *fra due gruppi* $\mathcal{L}$ *e* $\mathcal{H}$ *consiste nell'insieme delle coppie ordinate* $\mathcal{G} = \{(l_\alpha, h_\beta) : l_\alpha \in \mathcal{L}, h_\beta \in \mathcal{H}\}$ *ed è dotato della legge di composizione* $(l_\alpha \cdot_{\mathcal{L}} l_{\alpha'}, h_\beta \cdot_{\mathcal{H}} h_{\beta'})$. *Esso viene indicato come* $\mathcal{G} = \mathcal{L} \otimes \mathcal{H}$.

Non è difficile mostrare che esso è un gruppo. Ha senso definire anche il prodotto tensore di un gruppo per se stesso.

Definizione 6.5. *Il prodotto diretto di un gruppo X per se stesso N volte viene indicato solitamente con* X^n. *Così,* $\mathbb{R}^n$ *è il prodotto diretto di N copie del gruppo additivo e abeliano* $\mathbb{R}$.

6.2 Sottogruppi

Definizione 6.6. *Un* **sottogruppo** $\mathcal{L} \subseteq \mathcal{G}$ *è un sottoinsieme di un gruppo* $\mathcal{G}$ *che presenta anch'esso una struttura di gruppo rispetto alla stessa legge di composizione di* $\mathcal{G}$.

Esempio 6.9. Due esempi banali di sottogruppi sono $\mathcal{L}_0 = \{g_0\}$, che consiste del solo elemento neutro del gruppo $\mathcal{G}$, e $\mathcal{L}_{max} = \mathcal{G}$.

Esempio 6.10. L'insieme $Tr_x = \{T_\mathbf{v}, \mathbf{v} = (v, 0, 0)\}$ è il gruppo delle traslazioni lungo la direzione x e quindi è un sottogruppo del gruppo $\mathcal{G}_4$ precedentemente introdotto. Anche Tr_y e Tr_z, con analogo significato, sono dei sottogruppi di $\mathcal{G}_4$, così come in generale Tr_a, dove **a** è una direzione fissata. Analogamente, anche il sottoinsieme delle rotazioni in $\mathbb{R}^3$ attorno a una direzione fissata costituisce un sottogruppo abeliano di SO(3).

Definizione 6.7. *Dato un insieme di elementi* $\mathcal{V} = \{v_i\}$ *e un gruppo* $\mathcal{G}$, *l'insieme*

$$O(v) = \{gv \subset \mathcal{V} : g \in \mathcal{G}\} \tag{6.2}$$

è detto **orbita** *di* v.

Definizione 6.8. *Un gruppo è detto* **transitivo** *se* $\forall\, v_1, v_2 \in \mathcal{V}\ \ \exists g \in \mathcal{G} : v_1 = g\, v_2$, *ovvero se tutte le orbite relative ad ogni vettore coincidono.*

Esempio 6.11. $\mathcal{G}_v = \{g \in \mathcal{G} : gv = v\}$ è detto *gruppo di stabilità* di v. È un sottogruppo di $\mathcal{G}$.

Si può dimostrare il seguente teorema, che caratterizza i sottogruppi.

Teorema 6.1. *Sia* $\mathcal{G}$ *un gruppo e* $\mathcal{L} \subseteq \mathcal{G}$ *un suo sottoinsieme. Le seguenti affermazioni sono equivalenti:*

1. $\mathcal{L}$ *è un sottogruppo di* $\mathcal{G}$;
2. $\mathcal{L}$ *non è vuoto e* $\forall\, g_\alpha, g_\beta \in \mathcal{L}$ *risulta* $g_\alpha g_\beta \in \mathcal{L}$. *Inoltre,* $\forall\, g_\alpha \in \mathcal{L}$, $\exists g_\alpha^{-1} \in \mathcal{L}$;
3. $\mathcal{L}$ *non è vuoto e* $\forall\, g_\alpha, g_\beta \in \mathcal{L}$ *risulta* $g_\alpha g_\beta^{-1} \in \mathcal{L}$.

Dimostrazione. Dimostriamo il teorema provando le implicazioni in maniera ciclica.

($1 \Rightarrow 2$) se $\mathcal{L}$ è un sottogruppo, allora $g_0 \in \mathcal{L}$ e quindi l'insieme è non vuoto. Le altre due proprietà discendono banalmente dal fatto che $\mathcal{L}$ è un gruppo.

($2 \Rightarrow 3$) per ipotesi $\forall\, g_\beta \in \mathcal{L} \Rightarrow g_\beta^{-1} \in \mathcal{L}$. Poiché il prodotto di ogni coppia di elementi di $\mathcal{L}$ sta in $\mathcal{L}$ per ipotesi, allora anche $g_\alpha g_\beta^{-1} \in \mathcal{L}$.

($3 \Rightarrow 1$) $\mathcal{L}$ è non vuoto, quindi $\exists\, g_\alpha \in \mathcal{L}$. Si considerino allora due casi: se $g_0 \in \mathcal{L}$ ed è l'unico elemento, allora $\mathcal{L}$ è un sottogruppo banale; se $g_\alpha \in \mathcal{L}, \alpha \neq 0$, usando l'ipotesi con $\beta = \alpha$, si ottiene $g_\alpha g_\alpha^{-1} = g_0 \in \mathcal{L}$, e quindi $g_0 g_\alpha^{-1} = g_\alpha^{-1} \in \mathcal{L}$. Abbiamo quindi provato che l'inverso appartiene al sottogruppo $\mathcal{L}$. Poiché $(g_\beta^{-1})^{-1} = g_\beta$, allora $g_\alpha g_\beta = g_\alpha (g_\beta^{-1})^{-1} \in \mathcal{L}$. Quindi, $\mathcal{L}$ è un gruppo. □

Definizione 6.9. *L'insieme degli elementi del gruppo che commutano con ogni elemento del gruppo, ovvero* $Z(\mathcal{G}) = \{g_\alpha \in \mathcal{G} \mid g_\alpha g_\beta = g_\beta g_\alpha \,\forall\, g_\beta \in \mathcal{G}\}$, *è detto* centro *del gruppo.*

Non è difficile mostrare che il centro di un gruppo è un suo sottogruppo.

Proposizione 6.3. *$Z(\mathcal{G})$ è un sottogruppo di $\mathcal{G}$.*

Dimostrazione. Presi $g_\alpha, g_\beta \in Z(\mathcal{G})$, si ha $(g_\alpha g_\beta) g_\gamma = g_\alpha (g_\gamma g_\beta) = g_\gamma (g_\alpha g_\beta)$, ovvero $g_\alpha g_\beta \in Z(\mathcal{G})$. Inoltre, poiché $g_\alpha g_\beta = g_\beta g_\alpha$, moltiplicando a destra e a sinistra ambo i membri per g_α^{-1}, si ottiene $g_\alpha^{-1}(g_\alpha g_\beta) g_\alpha^{-1} = g_\alpha^{-1}(g_\beta g_\alpha) g_\alpha^{-1} \Rightarrow g_\beta g_\alpha^{-1} = g_\alpha^{-1} g_\beta$, ovvero $g_\alpha^{-1} \in Z(\mathcal{G})$. Questo, usando il Teorema 6.1, conclude la dimostrazione. □

All'interno di ogni gruppo si può introdurre la relazione

$$g_1 \sim g_2 \iff \exists\, g \in \mathcal{G} : g_2 = g\, g_1\, g^{-1}.$$

Non è difficile mostrare che essa è una relazione di equivalenza.

Proposizione 6.4. *$\sim$ è una relazione di equivalenza.*

Dimostrazione. Per dimostrare l'asserto, bisogna mostrare che valgono le tre proprietà

1. **riflessiva**: $g_1 \sim g_1 \iff g_1 = g\, g_1\, g^{-1}$, che risulta vera scegliendo $g = g_0$;
2. **simmetrica**: $g_1 \sim g_2 \iff g_1 = g\, g_2\, g^{-1}$, che risulta vera moltiplicando a destra per g^{-1} e a sinistra per g e ottenendo $g_2 = g^{-1} g_1\, g \Rightarrow g_2 \sim g_1$;
3. **transitiva**: se $g_1 \sim g_2 \iff g_1 = g\, g_2\, g^{-1}$ e $g_2 \sim g_3 \iff g_2 = h\, g_3\, h^{-1}$, allora scegliendo $k = gh$, si ha $g_1 = k\, g_3\, k^{-1} \iff g_1 \sim g_3$ e quindi risulta vera.

Questo completa la dimostrazione. □

Quindi $\mathcal{G}$ può essere quozientato tramite questa relazione di equivalenza. Non è difficile dimostrare la seguente proposizione.

Proposizione 6.5. *L'elemento g_0, elemento neutro del gruppo, è l'unico appartenente alla sua classe di equivalenza. Inoltre, un qualsiasi $h \in Z(\mathcal{G})$ è l'unico elemento appartenente alla sua classe di equivalenza.*

6.3 Gruppi di matrici

In teoria dei gruppi, i gruppi di matrici rivestono un ruolo di primo ordine in diverse applicazioni. Per questo motivo, nel seguente paragrafo verranno introdotti e discussi alcuni dei principali gruppi di matrici.

In generale, l'insieme $M_n(\mathbb{K}) = \{A \mid A \in \mathbb{K}^{n\times n}\}$ non è un gruppo in quanto non è detto che tutti i suoi elementi ammettano inverso. Come precedentemente argomentato in Esempio 6.6, per ottenere un gruppo da questo insieme basta restringersi al suo sottoinsieme con $\det A \neq 0$.

Definizione 6.10. *Si chiama* ***gruppo lineare generale*** *il gruppo delle matrici quadrate a determinante non nullo, ovvero*

$$\mathrm{GL}_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \,:\, \det A \neq 0\}. \tag{6.3}$$

Esempio 6.12. Mostriamo ora che $\mathrm{GL}_n(\mathbb{R})$ è un gruppo.

1. il prodotto tra matrici $n \times n$ è ancora una matrice $n \times n$;
2. la proprietà associativa è banale;
3. l'elemento neutro è la matrice identica $\mathbb{1}$;
4. $\exists\, A^{-1}$, poiché $\det A \neq 0$. Tale matrice inversa appartiene a $\mathrm{GL}_n(\mathbb{R})$ in quanto $\det A^{-1} = \dfrac{1}{\det A}$.

Presentiamo adesso alcuni sottogruppi notevoli del gruppo lineare generale.

Esempio 6.13. Il gruppo

$$\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) \,:\, \det A = 1\}$$

è detto **gruppo lineare speciale** e che esso sia un sottogruppo di $\mathrm{GL}_n(\mathbb{R})$ è evidente dal fatto che $AB^{-1} \in \mathrm{SL}_n(\mathbb{R})$ perché $\det(AB^{-1}) = \dfrac{\det A}{\det B} = 1$.

Esempio 6.14. Un altro sottogruppo notevole di $\mathrm{GL}_n(\mathbb{R})$ è il **gruppo ortogonale**

$$\mathrm{O}(n) = \{A \in M_n(\mathbb{R}) \,:\, AA^T = \mathbb{1}\}.$$

La condizione che $A \in \mathrm{GL}_n(\mathbb{R})$ non è esplicitamente richiesta nella definizione in quanto $AA^T = \mathbb{1}$ implica che $\det(AA^T) = \det A \det A^T = \det^2(A) = \det \mathbb{1} = 1 \Rightarrow \det A \neq 0$. Esso è un sottogruppo perché se $A, B \in \mathrm{O}(n)$, allora $(AB)(AB)^T = A(BB^T)A^T = A\mathbb{1}A^T = AA^T = \mathbb{1}$ e, inoltre, se $A \in \mathrm{O}(n)$, allora $A^{-1}(A^{-1})^T = (A^TA)^{-1} = \mathbb{1}^{-1} = \mathbb{1} \Rightarrow A^{-1} \in \mathrm{O}(n)$. La particolarità di tale gruppo è che esso mantiene inalterato il prodotto scalare in $\mathbb{R}^n$, ovvero se $A \in \mathrm{O}(n)$, allora $\langle Aa, Ab\rangle = \langle a, b\rangle$. In $\mathrm{O}(n)$ stanno tutte le rotazioni e le riflessioni in $\mathbb{R}^n$, che sono rappresentate da matrici ortogonali. Invece, nel **gruppo ortogonale speciale** $\mathrm{SO}(n) = \{A \in \mathrm{O}(n) \,:\, \det A = 1\}$ stanno solo le rotazioni, nel caso in cui n sia dispari, altrimenti se n è pari vi stanno anche le riflessioni.

Esempio 6.15. Un gruppo notevole è il **gruppo simplettico**. Data una matrice $J \in M_{2n}(\mathbb{R})$ tale che

$$J = \begin{pmatrix} 0_n & \mathbb{1}_n \\ -\mathbb{1}_n & 0_n \end{pmatrix} \tag{6.4}$$

il gruppo simplettico è $\text{sp}(2n, \mathbb{R}) := \{A \in M_{2n}(\mathbb{R}) \, : \, A^T J A = J\}$.

Esempio 6.16. I gruppi finora presentati hanno in comune la proprietà di essere sottoinsiemi del gruppo $M_n(\mathbb{R})$, contenente matrici a entrate reali. I gruppi precedenti possono essere estesi anche al caso in cui l'insieme di lavoro sia $M_n(\mathbb{C})$, insieme delle matrici quadrate a entrate complesse. Si possono però in questo contesto introdurre ulteriori gruppi notevoli. Uno di questi è il **gruppo unitario** $\text{U}(n) = \{A \in M_n(\mathbb{C}) \, : \, AA^\dagger = \mathbb{1}\}$. Analogamente al gruppo ortogonale, anche gli elementi del gruppo unitario conservano il prodotto scalare in $\mathbb{C}^n$. Inoltre, la proprietà con cui è definito tale gruppo implica che $|\det A| = 1$. Per selezionare soltanto le matrici con determinante esattamente uguale a 1, si introduce il **gruppo unitario speciale** $\text{SU}(n) = \{A \in \text{U}(n) \, : \, \det A = 1\}$.

6.4 Rappresentazione di un gruppo

Definizione 6.11. *Una mappa $f : \mathcal{G} \to \mathcal{S}$, con $\mathcal{G}, \mathcal{S}$ gruppi, è detta un* **omomorfismo** *se conserva le operazioni dei gruppi, ovvero se risulta che*

$$g_\alpha g_\beta = g_\gamma \Rightarrow f(g_\alpha g_\beta) = f(g_\alpha) f(g_\beta) = f(g_\gamma) \tag{6.5}$$

Se f è invertibile, essa è detto un **isomorfismo**. *Se $\mathcal{G} = \mathcal{S}$, allora f è detto* **endomorfismo**. *Se f è un endomorfismo invertibile, allora è detto un* **automorfismo**.

Proposizione 6.6. *Se f è un omomorfismo, detti g_0 e s_0 elementi neutri dei due gruppi, allora $s_0 = f(g_0)$.*

Dimostrazione. Siano $g \in \mathcal{G}$ e $s = f(g) \in \mathcal{S}$. Allora $s = f(g) = f(g_0 g) = f(g_0) f(g) = f(g_0) s$, da cui segue la tesi. □

Esempio 6.17. Il determinante è un esempio di omomorfismo, in quanto è una mappa $\det : \text{GL}_n(\mathbb{C}) \to \mathbb{C} \setminus \{0\}$ tale che $\det AB = \det A \det B$.

Il fatto che un'applicazione sia un omomorfismo ha una serie di conseguenze interessanti.

Teorema 6.2. *Sia f un omomorfismo. Allora*

1. $(f(g))^{-1} = f(g^{-1}) \; \forall \, g \in \mathcal{G}$;
2. *Se $\mathcal{G}_1$è un sottogruppo di $\mathcal{G}$, allora la sua immagine mediante f è un sottogruppo di $\mathcal{S}$;*
3. *L'insieme $\text{Im}(f) = \{f(g) \, : \, g \in \mathcal{G}\}$ è un sottogruppo di $\mathcal{S}$;*
4. *L'insieme $\ker(f) = \{g \in \mathcal{G} \, : \, f(g) = s_0\}$ è un sottogruppo di $\mathcal{G}$.*

Dimostrazione.
1. $s_0 = f(g_0) = f(gg^{-1}) = f(g) f(g^{-1}) \Rightarrow (f(g))^{-1} s_0 = f(g^{-1})$ che implica la tesi;
2. Sia $\mathcal{S}_1 = \{f(g) \, : \, g \in \mathcal{G}_1\} \neq \emptyset$ perché $g_0 \in \mathcal{G}_1$ e quindi $s_0 = f(g_0) \in \mathcal{S}_1$. Se $s_\alpha, s_\beta \in \mathcal{S}_1 \iff \exists \, g_\alpha, g_\beta \, : \, s_{\alpha,\beta} = f(g_{\alpha,\beta})$. Allora $s_\alpha s_\beta^{-1} = f(g_\alpha)(f(g_\beta))^{-1} = f(g_\alpha g_\beta^{-1}) \in \mathcal{S}_1$, in quanto $g_\alpha g_\beta^{-1} \in \mathcal{G}_1$;
3. Poiché $\mathcal{G}$ è un sottogruppo banale di se stesso, dalla 2. segue la tesi;

4. $\ker f$ non è vuoto, in quanto $g_0 \in \ker f$. Se $g_\alpha, g_\beta \in \ker f \Rightarrow f(g_\alpha g_\beta^{-1}) = f(g_\alpha)f(g_\beta^{-1}) = f(g_\alpha)(f(g_\beta))^{-1} = s_0 s_0^{-1} = s_0$. Quindi $\ker f$ è un sottogruppo di $\mathcal{G}$.

□

A questo punto, si può introdurre l'utile concetto di rappresentazione di un gruppo.

Definizione 6.12. *Una* **rappresentazione** *di un gruppo $\mathcal{G}$ è un omomorfismo $\mathcal{R}$ da $\mathcal{G}$ in $\mathrm{GL}_n(\mathbb{C})$. Essa è detta* **fedele** *se l'omomorfismo è iniettivo.*

Definizione 6.13. *Due rappresentazioni $\mathcal{R}_1$ e $\mathcal{R}_2$ dello stesso gruppo $\mathcal{G}$ sono dette* **equivalenti** *se $\mathcal{R}_1(g) = V\mathcal{R}_2(g)V^{-1}\ \forall\, g \in \mathcal{G}$.*

Definizione 6.14. *$\mathcal{R}$ è una rappresentazione* unitaria *se la matrice $\mathcal{R}(g)$ è unitaria $\forall\, g \in \mathcal{G}$.*

Esempio 6.18. Si consideri il gruppo discreto $\mathcal{G} = \{R_0, R_{\pi/2}, R_\pi, R_{3\pi/2}\}$, dove R_α è una rotazione di un angolo α. Una possibile rappresentazione del gruppo è $\mathcal{R}_1$ tale che

$$u_0 = -u_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad u_1 = -u_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Un'altra rappresentazione del gruppo è tale che

$$v_0 = \mathbb{1}_4 \quad , \quad v_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad , \quad v_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad , \quad v_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Le due rappresentazioni non sono però equivalenti e questo è evidente dal fatto che le matrici hanno dimensioni diverse e quindi agiscono su spazi diversi.

6.5 Gruppi di Lie

Tra i gruppi, un ruolo di particolare importanza è rivestito dai **gruppi di Lie**. Essi trovano una grande applicazione in meccanica, dove vengono usati per modellizzare le trasformazioni che un sistema può subire e per studiarne le simmetrie. Essi sono un particolare caso dei più generali gruppi dipendenti da dei parametri.
In generale, gli elementi di un gruppo possono dipende da uno o più parametri esterni, che possono variare discretamente o con continuità.

Esempio 6.19. Si consideri l'operatore di traslazione $T_a f(x) = f(x+a)$. L'insieme $\mathcal{G}_t = \{T_a\}$ è un gruppo abeliano rispetto al prodotto usuale di operatori. Infatti, $T_0 = \mathbb{1}$ è l'elemento neutro del gruppo, T_{-a} è l'inverso di T_a e il prodotto di due traslazioni è ancora una traslazione $T_a T_b = T_{a+b}$. Inoltre, non è difficile convincersi che $T_a T_b = T_b T_a$. Si noti che $\lim_{a,0} T_a f(x) = f(x) = \mathbb{1} f(x)$. Le proprietà enunciate continuano a valere anche considerando traslazioni tridimensionali.

Esempio 6.20. Nello spazio tridimensionale ordinario, si consideri l'operatore R_θ di rotazione intorno all'asse z. L'insieme $\mathcal{G}_r(1) = \{R_\theta\}$ forma un gruppo non abeliano abeliano rispetto al prodotto

usuale di operatori. Una rappresentazione di tale gruppo può essere fornita dalla matrice

$$R_\theta \mathbf{r} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{r} \qquad \theta \in [0, 2\pi]$$

Anche in questo caso, vale la proprietà che $\lim_{\theta,0} R_\theta = \mathbb{1}$. È evidente che deve esistere una topologia per poter effettuare questa operazione.

Definizione 6.15. *Un* **gruppo di Lie** *è un gruppo, i cui elementi dipendono da un set di parametri* $\mathbf{a} = (a_1, a_2, \ldots, a_m), m \geq 1$, *dotato di una norma* $\|\cdot\|$ *tale che la mappa che associa ad ogni coppia di elementi il loro prodotto e la mappa che associa ad ogni elemento il suo inverso siano ambedue differenziabili.*

Una situazione abbastanza generale è quella dei **gruppi di Lie di matrici**, in cui gli elementi del gruppo sono matrici quadrate $\{\mathcal{M}(\mathbf{a})_{ij}\}$. In questo caso, si intendono la derivata e il differenziale della matrice come la derivata e il differenziale delle singole entrate, che sono in generale funzioni del set di parametri **a**.

Definizione 6.16. *Un gruppo di Lie si dice* **compatto** *se il set di parametri varia in un insieme compatto.*

Esempio 6.21. Facendo riferimento al gruppo delle rotazioni introdotto in Esempio 6.20, esso è un gruppo di Lie compatto in quanto θ varia in un intervallo chiuso e limitato. Il gruppo introdotto in Esempio 6.19 è un gruppo di Lie ma non compatto.

Definizione 6.17. *Un gruppo di Lie è detto* connesso *se presi* $g_1, g_2 \in \mathcal{G}\ \exists g(a) \in \mathcal{G}$ *dipendente da a con continuità tale che* $g_1 = g(a_1)$ *e* $g_2 = g(a_2)$ *per qualche valore di* a_1, a_2.

Esempio 6.22. Il gruppo delle riflessioni non è continuo. Allora, non può essere connesso.

6.6 Algebra di Lie

Definizione 6.18. *Un'***algebra di Lie** *è uno SVL V su campo* $\mathbb{K}$ *dotato di un'operazione* $[\![\cdot,\cdot]\!]$ *detta* **parentesi di Lie** *tale che* $[\![\cdot,\cdot]\!] : V \times V \to V$ *e inoltre*

1. $[\![\cdot,\cdot]\!]$ *è bilineare, ovvero lineare in entrambe le componenti;*
2. $[\![x,y]\!] = -[\![y,x]\!]\ \forall x, y \in V$ *(**antisimmetrica**).*
3. *vale l'**identità di Jacobi***

$$[\![x, [\![y,z]\!]]\!] + [\![y, [\![z,x]\!]]\!] + [\![z, [\![x,y]\!]]\!] = 0\ \forall\, x, y, z \in V. \tag{6.6}$$

Esempio 6.23. Due esempi semplici di algebre di Lie sono lo spazio $\mathbb{R}^3$, dotato del prodotto vettoriale ordinario, e lo spazio $B(H)$ degli operatori limitati di H in H ($B(H) \in \mathrm{End}(H)$), dotato del commutatore tra operatori.

Proposizione 6.7. *Sia* $\mathcal{M}$ *un gruppo di Lie di matrici. L'insieme* $\mathfrak{M} = \{X \in M_n(\mathbb{C}) : e^{tX} \in \mathcal{M}\ \forall t \in \mathbb{R}\}$ *è l'algebra di Lie associata a* $\mathcal{M}$ *con parentesi di Lie definita dal commutatore tra operatori.*

Definizione 6.19. *I* ***generatori del gruppo*** *sono gli elementi* $\{B_{\mathcal{M}}\} = \{T_i\}_{i=1}^{d_{\mathcal{M}}}$ *di una base di* $\mathfrak{M}$*, che è uno spazio vettoriale lineare di dimensione* $d_{\mathcal{M}}$*.*

Quindi, ogni elemento di $\mathfrak{M}$ può essere espanso in serie, fissato che sia un insieme di generatori del gruppo. Il commutatore tra due elementi di una base di $\mathfrak{M}$ può essere scritto come $[T_i, T_j] = \sum_{k=1}^{d_{\mathcal{M}}} c_{ijk} T_k$. Le costanti c_{ijk} sono dette *costanti di struttura* e sono fortemente connessi alle specifiche del gruppo. Data l'antisimmetria del commutatore, è evidente che $c_{ijk} = -c_{jik}$. A questo punto, non è difficile mostrare che il commutatore definisce effettivamente una parentesi di Lie per l'insieme $\mathfrak{M}$. Infatti, dati $X, Y \in \mathfrak{M}$, si ha

$$\boxed{[X, Y] = \sum_{i,j=1}^{d_{\mathcal{M}}} x_i y_i [T_i, T_j] = \sum_{i,j,k=1}^{d_{\mathcal{M}}} x_i y_i c_{ijk} T_k = \sum_{k=1}^{d_{\mathcal{M}}} \alpha_k T_k \in \mathfrak{M}} \tag{6.7}$$

L'identità di Jacobi può essere riscritta in termini delle costanti di struttura ottenendo

$$\sum_{k=1}^{d_{\mathcal{M}}} (c_{jlk} c_{ikm} + c_{lik} c_{jkm} + c_{ijk} c_{lkm}) = 0 \tag{6.8}$$

Nelle applicazioni fisiche, si è soliti definire l'algebra associata a un gruppo come l'insieme delle matrici autoaggiunte tali che $e^{itX} \in \mathcal{M}$. In questa maniera, gli elementi dell'algebra sono operatori unitari e quindi limitati. Se $M(a)$ è il generico elemento del gruppo $\mathcal{M}$, si ha $M(a) = \exp\left\{ i \sum_{j=1}^{d_{\mathcal{M}}} a_j T_j \right\}$. Perciò i generatori del gruppo possono essere ottenuti derivando come

$$\boxed{T_j = -i \left. \frac{\partial M(a)}{\partial a_j} \right|_{a=0}} \tag{6.9}$$

6.6.1 Esempi notevoli di algebre di Lie

In questo paragrafo sono raccolti una serie di esempi più o meno importanti di algebre di Lie.

Esempio 6.24. Si consideri il gruppo $\mathrm{U}(1) = \{e^{i\alpha}, \alpha \in \mathbb{R}\}$. Esso è un algebra di Lie e ha un unico generatore, ovvero la matrice identica $\mathbb{1}$.

Esempio 6.25. Si consideri il gruppo delle rotazioni nello spazio tridimensionale SO (3). Fissati che siano gli assi di rotazione, nell'ordine x, y, z, le rotazioni intorno a tali assi sono rappresentate dalle seguenti matrici:

$$R_x(\theta_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & \sin\theta_x \\ 0 & -\sin\theta_x & \cos\theta_x \end{pmatrix}$$

$$R_y(\theta_y) = \begin{pmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{pmatrix}$$

$$R_z(\theta_z) = \begin{pmatrix} \cos\theta_z & \sin\theta_z & 0 \\ -\sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

da cui si ottengono, derivando come da Equazione 6.9, i generatori

$$T_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad T_y = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \qquad T_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Quindi si ottiene che i generatori sono hermitiani e inoltre $[T_i, T_j] = i\sum_k \epsilon_{ijk}T_k$, dove ϵ_{ijk} è il tensore completamente antisimmetrico.

Esempio 6.26. Dal gruppo SU(2) si ottengono le **matrici di Pauli** come generatori e le stesse costanti di struttura.

Esempio 6.27. Il **gruppo di Galilei** è il gruppo a 4 parametri $(\mathbf{R}, \mathbf{v}, \mathbf{y}, \tau)$, rispettivamente relativi a una rotazione, un cambio di velocità, una traslazione spaziale e una traslazione temporale, composti dalle trasformazioni di Galileo

$$g_{\mathbf{R},\mathbf{v},\mathbf{y},\tau}(\mathbf{r}, t) = (\mathbf{R}\,\mathbf{r} + \mathbf{v}t + \mathbf{y}, t + \tau)$$

essendo un gruppo a 10 parametri totali, ha altrettanti generatori. L'identità del gruppo è l'elemento $g_{\mathbb{1},\mathbf{0},\mathbf{0},0}$.

Esempio 6.28. Il **gruppo di Lorentz**, a volte indicato con O(3,1) è il gruppo delle matrici a entrate reali che conservano la metrica di Minkowski $\eta = \mathrm{diag}(+1, -1, -1, -1)$. Una matrice Λ quindi appartiene a $\mathcal{G}_{Lor}$ se $\Lambda^T \eta \Lambda = \eta$. Questo implica che $\det\Lambda = \pm 1$.

Esempio 6.29. Il **gruppo di Poincaré** $\mathcal{G}_{Poin}$ è il gruppo i cui elementi rappresentano la composizione di una trasformazione di Lorentz e di una traslazione del quadrivettore

$$\mathcal{G}_{Poin} := \{(\Lambda, a) : \Lambda \in \mathcal{G}_{Lor}, a \in \mathbb{R}^4\} = (\Lambda, a)\mathbf{r} = \Lambda\mathbf{r} + \mathbf{a}$$

Le legge di composizione degli elementi del gruppo è $(\Lambda', a')(\Lambda, a) = (\Lambda'\Lambda, \Lambda' a + a)$. L'inverso di un elemento del gruppo è $(\Lambda, a)^{-1} = (\Lambda^{-1}, -\Lambda^{-1}a)$ e l'elemento neutro è $(\mathbb{1}_4, 0)$.

6.6.2 Gruppo delle simmetrie

Definizione 6.20. *Dato un osservabile $S \in B(H)$ di un sistema quantistico $\mathcal{S}$ caratterizzato da un operatore hamiltoniano H autoaggiunto, esso è una* simmetria del sistema *se $[H, S] = 0$ e se S è invertibile.*

Questa definizione è giustificata perché, lavorando in rappresentazione di Heisenberg e con hamiltoniana non dipendente esplicitamente dal tempo, si ha che

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = i\,[H, X(t)] = i\,\mathrm{e}^{iHt}\,[H, X]\,\mathrm{e}^{-iHt} \tag{6.10}$$

L'invertibilità è necessaria in quanto se un sistema è invariante per una certa trasformazione, è lecito aspettarsi che sia invariante per la trasformazione inversa. La possibilità di scrivere

un'espressione del tipo (6.10) ha una conseguenza estremamente interessante, sintetizzata nel seguente teorema.

Teorema 6.3 (Noether). *Sia S una simmetria. Allora, essa è una costante del moto, ovvero $\dot{S} = 0$*

Proposizione 6.8. *Sia $\mathbb{D} = \{S : S \text{ è una simmetria}\}$. L'insieme $\mathbb{D}$ forma un gruppo rispetto alla moltiplicazione di operatori.*

Dimostrazione. $\mathbb{1} \in \mathbb{D}$, poiché commuta con l'hamiltoniano ed è invertibile. Prese due simmetrie S_1, S_2, si ha che $(S_1 S_2)^{-1} = S_2^{-1} S_1^{-1}$ e inoltre $[H, S_1 S_2] = [H, S_1] S_2 + S_1 [H, S_2] = 0$, quindi $\mathbb{S}$ è chiuso rispetto alla moltiplicazione di operatori. La proprietà associativa del gruppo è banale. Infine, presa una simmetria S, sicuramente S^{-1} esiste e inoltre vale che $[H, S^{-1}] = 0$ perché, se $[H, S] = 0$ allora $H S = S H$ e quindi moltiplicando a destra e a sinistra per S^{-1} si ottiene la tesi. □

Appendice A

Richiami di calcolo complesso

Il calcolo differenziale e integrale per funzioni di singola variabile complessa è indispensabile in molte sezioni di questo libro. In questa appendice, riportiamo e giustifichiamo brevemente alcuni risultati fondamentali. Per ulteriori approfondimenti, si consiglia la lettura di *Fundamentals of Complex Analysis with applications to engineering, science, and mathematics, 2013 Pearson Ed. Ltd., 3a ed.* di Saff, Snider.

A.1 Funzioni analitiche

Una funzione di variabile complessa $f(z)$ può essere scritta come

$$f(z) = u(x,y) + iv(x,y) \quad \text{con} \quad z = x + iy. \tag{A.1}$$

Una funzione di variabile complessa della forma di è detta *analitica* se obbedisce alle seguenti condizioni, dette di **Cauchy-Riemann** (*CRCs*):

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \wedge \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{A.2}$$

Elenchiamo adesso alcuni teoremi sulle funzioni analitiche.

Teorema A.1. *Sia $f(z)$ una funzione analitica. Le seguenti condizioni sono equivalenti*

1. *Valgono le condizioni di Cauchy-Riemann;*
2. $\dfrac{\partial f}{\partial x} + i\dfrac{\partial f}{\partial y} = 0$;
3. $\exists \lim_{z,z_0} \dfrac{f(z) - f(z_0)}{z - z_0} =: f'(z)$, *ovvero la funzione è derivabile.*

Basta dimostrare l'esistenza della derivata prima per essere certi che la funzione sia infinitamente derivabile, così come asserito nel seguente teorema.

Teorema A.2. *Data $f(z)$, se per un certo $k \in \mathbb{N}\ \exists\ f^{(k)}(z)$, ovvero se $f(z) \in C^k(\mathcal{E})$ segue che $f(z) \in C^\infty(\mathcal{E})\ \forall\ \mathcal{E} \subseteq \mathbb{C}$ aperto.*

A.2 Risultati notevoli sul calcolo integrale

Di fatto, una funzione di singola variabile complessa non è altro che una funzione di due variabili reale, ovvero la parte reale e immaginaria della variabile z. Per questo motivo, in analogia con gli

integrali curvilinei, è possibile definire l'integrale di una funzione di variabile complessa lungo una curva γ semplice a tratti, percorsa da z_0 a z_f, rispettivamente punto iniziale e finale della curva, come

$$\int_\gamma f(z)\mathrm{d}z := \int_\gamma (u\mathrm{d}x - v\mathrm{d}y) + i\int_\gamma (u\mathrm{d}y + v\mathrm{d}x). \tag{A.3}$$

L'estrema regolarità delle funzioni analitiche si manifesta anche nel risultato dell'integrazione curvilinea. Infatti si può dimostrare che se i cammini di integrazioni sono *equivalenti fra loro*, allora il risultato dell'integrale sarà lo stesso. Questo risultato è riassunto dal seguente teorema.

Teorema A.3. *Se $f(z)$ è analitica, dette γ_1, γ_2 due curve aventi in comune il punto iniziale e quello finale e omotope, allora*

$$\int_{\gamma_1} f(z)\mathrm{d}z = \int_{\gamma_2} f(z)\mathrm{d}z$$

Ciò significa che, se posso *deformare con continuità* γ_1 riuscendo a sovrapporla a γ_2, senza passare per *punti critici*, allora i due integrali si equivalgono. L'operazione, ad esempio, viene impedita dalla presenza di singolarità nell'area compresa tra le due curve.

Di particolare interesse sono gli integrali in cui γ è una curva chiusa. In questo caso, la curva partiziona il piano complesso in due regioni, una interna e una esterna alla curva γ. Indichiamo con γ^o il sottoinsieme limitato di $\mathbb{C}$ delimitato da γ. Un risultato importante, attribuito a Cauchy, è riassunto nella seguente proposizione.

Proposizione A.1. *Data una funzione $f(z)$ analitica e una curva chiusa γ, se la funzione non presenta singolarità* [a] *nella parte di piano interna alla curva γ, allora* $\oint_\gamma f(z)\mathrm{d}z = 0$.

[a] Si dice *punto singolare*, o *singolarità*, un punto in cui la funzione non è definita.

Questo risultato può essere ulteriormente generalizzato, ottenendo il seguente teorema.

Teorema A.4 (di Cauchy). *Se f è analitica, γ è chiusa e γ^o non contiene singolarità di f, allora vale che*

$$\frac{1}{2\pi i}\oint_\gamma \frac{f(z)}{z - z_0}\mathrm{d}z = \begin{cases} 0 & se\ z_0 \notin \gamma^o \\ f(z) & se\ z_0 \in \gamma^o \end{cases}$$

Questo comporta la conoscenza dei valori che la funzione f assume sulla curva γ permette di ricostruire i valori di f all'interno di γ. La stessa formula vale per le derivate di ordine superiore. Infatti, si trova facilmente che

$$\boxed{f^{(k)}(z) = \frac{k!}{2\pi i}\oint_\gamma \frac{f(z)}{(z - z_0)^k}\mathrm{d}z \quad \forall z_0 \in \gamma^o.} \tag{A.4}$$

A.2.1 Il teorema dei residui

Introduciamo adesso il concetto di residuo di una funzione complessa in un punto del piano complesso.

Definizione A.1 (Teorema dei residui). *Siano $f(z)$ una funzione analitica e z_0 un punto del piano complesso. Scelta una curva γ chiusa e semplice a tratti tale che $z_0 \in \gamma^o$, si dice* **residuo** *di $f(z)$ in z_0 il valore*

$$\text{Res}\{f(z)\}\Big|_{z_0} = \frac{1}{2\pi i} \oint_{\gamma} f(z)\mathrm{d}z. \tag{A.5}$$

Il valore del residuo non dipende dalla scelta della curva γ, a patto che il punto z_0 sia contenuto nella parte di piano γ^o delimitata dalla curva. Se inoltre γ è orientata positivamente, si può scambiare senza indugio il simbolo di integrale $\oint$ con $\oint$. Quando ciò non accade, è necessario porre attenzione ai segni che compaiono.

Singolarità isolate

Sia z_0 un punto contenuto in un aperto $I \subseteq \mathbb{C}$. Si dice che

Definizione A.2. *$f(z)$ possiede una* **singolarità isolata** *in z_0 se esiste un intorno I di z_0 per cui la funzione è analitica in $I \setminus \{z_0\}$.*

Quindi la funzione non è definita in z_0, mentre lo è, ed è ivi infinitamente derivabile, in ogni altro punto sufficientemente vicino. Una singolarità isolata è detta

1. **eliminabile** se esiste il
$$\lim_{z,z_0} f(z) = l \in \mathbb{C};$$
2. **polo** di ordine n se esiste un $n \in \mathbb{N}$ tale che esista il
$$\lim_{z,z_0} (z - z_0)^n f(z) = l \in \mathbb{C};$$
3. **essenziale** se non rientra nei casi precedenti.

Calcolo dei residui

Il calcolo dei residui è particolarmente semplice nei primi due casi. Ad esempio, nel caso di una singolarità eliminabile, il residuo è pari a 0. Si può dimostrare che il residuo corrisponde al coefficiente di ordine -1 dello sviluppo in serie di Laurent della funzione $f(z)$.

Teorema A.5 (calcolo dei residui). *Se $f(z)$ presenta un polo di ordine n in z, allora*

$$\text{Res}\{f(z_0\} \equiv a_{-1} = \frac{1}{(n-1)!} \lim_{z,z_0} \frac{\mathrm{d}^{n-1}}{\mathrm{d}z^{n-1}}[(z - z_0)^n \cdot f(z)]. \tag{A.6}$$

La formula si semplifica molto nel caso in cui $n = 1$. Come si può immaginare, questa formula è indispensabile il calcolo del residuo.

Esempio A.1. Mostriamo adesso un esempio generale dell'utilizzo dei residui nel calcolo integrale. Si supponga di voler calcolare il valore di $I = \int_{\mathbb{R}} f(x)\,\mathrm{d}x$. Estendendo il dominio di definizione

di f al piano complesso, si può scrivere che

$$I = \int_{\mathbb{R}} f(z)\,\mathrm{d}z = \lim_{R,+\infty} \int_{-R}^{R} f(z)\,\mathrm{d}z = \lim_{R,+\infty} \left[\oint_{\Gamma_R} f(z)\,\mathrm{d}z - \int_{C_R(0,+)} f(z)\,\mathrm{d}z \right] \tag{A.7}$$

dove $C_R(0,+)$ è la semiconferenza, nel semipiano delle $\Im(z) > 0$, di raggio R e centrata in 0 e Γ_R è il percorso totale formato da $C_R(0,+)$ e dal diametro $[-R,+R]$. Notiamo adesso che se l'integrale sulla semiconferenza andasse a zero nel limite $R \to +\infty$, allora I si ridurrebbe all'integrale della funzione sul semipiano complesso con $\Im(z) > 0$. Questo permette l'applicazione del **teorema dei residui**, ovvero

$$I = \lim_{R,+\infty} \oint_{\Gamma_R} f(z)\,\mathrm{d}z = 2\pi i \sum_k \mathrm{Res}\{f\}\Big|_{z_k} \quad \text{con} \quad \Im(z_k) > 0. \tag{A.8}$$

ove il simbolo $\oint$ sostituisce $\oint$ poiché è stata definita in precedenza l'orientazione positiva della curva.

Il **lemma di Jordan**, nelle sue due formulazioni, fornisce delle condizioni sufficienti affinché l'integrale sulla semiconferenza tenda a zero.

Lemma A.1 (di Jordan, 1). *Sia $f(z)$ analitica e t.c. $\exists\, C, \epsilon > 0 : |f(z)| \leq \frac{C}{|z|^{1+\epsilon}}$ $\forall z \in \{\mathbb{C} : \Im(z) > 0\}$. Allora*

$$\lim_{R,+\infty} \int_{C_R(0,+)} f(z)\,\mathrm{d}z = 0. \tag{A.9}$$

Lemma A.2 (di Jordan, 2). *Sia $f(z)$ analitica tale che $|f(Re^{i\theta})| \leq \epsilon(R), \theta \in [0,\pi]$, con $\epsilon(x)$ funzione positiva e infinitesima per $R \to +\infty$. Allora, se $\alpha > 0$, si ha che*

$$\lim_{R,+\infty} \int_{C_R(0,+)} e^{i\alpha z} f(z)\,\mathrm{d}z = 0. \tag{A.10}$$

Nella pratica, vengono utilizzate anche versioni *alternative* di questo lemma in cui ci si trova, ad esempio, in $\Im(z) < 0$ o la curva viene percorsa in senso orario.

Appendice B

Misura e integrazione di Lebesgue

Per introdurre gli spazi di Hilbert, è necessario estendere la classe delle funzioni *integrabili*, superando con ciò l'integrale di Riemann[1]. Questa necessità è dettata dal fatto che l'integrale definito secondo Riemann riesce a integrare soltanto funzioni abbastanza *regolari*[2]. Nella pratica, però, si può essere nelle condizioni di dover modellizzare fenomeni caratterizzati da molto rumore, che analiticamente si traducono in funzioni ricche di discontinuità, e si rischia di trovarsi a lavorare con funzioni non Riemann-integrabili o comunque il cui integrale risulti difficile da calcolare.

Questa non è l'unica motivazione che spinge a forzare i limiti dell'integrazione. L'integrale di Riemann mostra infatti cattive proprietà di convergenza in alcune situazioni, come ed esempio quando si vuole capire se vale che

$$\int_D f(x)\ \mathrm{d}x = \int_D \lim_{n,\infty} f_n(x)\,\mathrm{d}x \stackrel{?}{=} \lim_{n,\infty} \int_D f_n(x)\,\mathrm{d}x\,. \tag{B.1}$$

Nel caso dell'integrale di Riemann, si conclude che una condizione sufficiente è chiedere la **convergenza uniforme** della successione di funzioni nel dominio di integrazione, essendo questo un intervallo finito. Tale condizione non può essere ulteriormente alleggerita e quindi di fatto limita di molto l'utilizzo di questa formula. Sarebbe infatti comodo, ad esempio, poter integrare gli sviluppi in serie di potenze delle funzioni più complicate da $\mathcal{R}$-integrare. Limitazioni simili sono presenti anche nel caso del passaggio al limite sotto il segno di derivata.

Per superare queste difficoltà, occorre introdurre l'*integrale di Lebesgue*. La sua introduzione, però, non può prescindere dal problema della **misurabilità** di un insieme nel senso di Lebesgue.

B.1 Misura di Lebesgue

L'ambiente di lavoro in cui viene sviluppata la teoria è $\mathbb{R}$. Si consideri un qualsiasi sottoinsieme $I \subseteq \mathbb{R}$. Se I fosse un intervallo chiuso e limitato, allora la definizione della sua misura sarebbe banale. Si pone infatti

$$\mu([a,b]) = b - a \tag{B.2}$$

[1]Molto interessante è la storia del calcolo integrale. Indipendentemente, tre famosi scienziati formularono il teorema fondamentale del calcolo, che associa il calcolo di un'area al calcolo di una primitiva. Tuttavia, quasi due secoli passarono prima che questi concetti venissero formalizzati da R., il quale in compenso ebbe una vita breve e morì di tubercolosi durante un viaggio in Italia. (*Bernhard Georg Friedrich Riemann*, ★1826, ≈(?), †1866, ⚱Verbania [IT], ⚭*Elise Koch.*) Inoltre, la sua tomba sarebbe stata distrutta da alcuni lavori di ristrutturazione.

[2]Quanto regolari è chiarito dal **Teorema di Vitali-Lebesgue**: una funzione limitata definita su un intervallo limitato è Riemann-integrabile se, e soltanto se, è continua quasi ovunque.

Tale definizione può essere estesa senza problema anche alle unioni finite di intervalli a due a due disgiunti. È interessante chiedersi come definire la misura di oggetti come

$$I = [3,7] \cup \{2, \pi, 8\} \tag{B.3}$$

che chiaramente non sono intervalli. Si definisce **ricoprimento esterno** un'unione di intervalli a due a due disgiunti contenente (al più uguale a) I. Così ad esempio

$$R^1_{ext}(I) = [0,10]$$
$$R^2_{ext}(I) = [2,8]$$
$$R^3_{ext}(I) = [2,2.5] \cup [3,7] \cup [7.5,8]$$

sono tutti ricoprimenti esterni ammissibili per (B.3). A questo punto, si pone la seguente

Definizione B.1. *Si definisce* ***misura esterna di I*** *come* $\mu_{ext}(I) = \inf\{\mu(R_{ext}(I))\}$.

Questa definizione è ben posta in quanto, essendo μ una funzione[3] non negativa, l'insieme

$$\{\mu(R_{ext}(I)) : R_{ext}(I) \in \mathcal{R}(I)\} \tag{B.4}$$

è sicuramente limitato inferiormente e quindi, per completezza di $\mathbb{R}$, ammette estremo inferiore. Analogamente, si definisce il concetto di **ricoprimento interno** e quindi

Definizione B.2. *Si definisce* ***misura interna di I*** *come* $\mu_{int}(I) = \sup\{\mu(R_{int}(I))\}$.

Vale sicuramente che

$$\mu_{int}(I) \leq \mu_{ext}(I) \tag{B.5}$$

Date queste definizioni, si può finalmente concludere che

Definizione B.3. *Si dice che un insieme è* **Lebesgue-misurabile** *se* $\mu_{int}(I) = \mu_{ext}(I) = \mu(I)$.

B.1.1 Misure di insiemi notevoli

1. **singleton**: tutti gli insiemi che constano di un solo punto sono $\mathscr{L}$-misurabili e hanno misura nulla. Questo si può dimostrare facilmente considerando che, essendo vuoto l'insieme dei ricoprimenti interni, sicuramente un singleton ha misura interna nulla, mentre la misura dei ricoprimenti esterni può essere resa arbitrariamente piccola;

2. insiemi formati da un **numero finito** di punti sono $\mathscr{L}$-misurabili e hanno misura nulla;

3. $\mathbb{N}$ è misurabile e ha misura nulla. Sicuramente infatti $\mu_{int} = 0$ perché l'unico ricoprimento ammissibile è il vuoto, che ha misura nulla. Si consideri ora un ricoprimento esterno formato da intorni dei punti di $\mathbb{N}$, ognuno di ampiezza $\epsilon/2^n$. In questo modo

$$\forall \epsilon > 0 \Rightarrow 0 \leq \mu_{ext}(\mathbb{N}) < \sum_{n=0}^{\infty} \frac{\epsilon}{2^n} = 2\epsilon \tag{B.6}$$

   da cui segue immediatamente che $\mu_{ext} = 0$ e quindi la tesi;

4. con un procedimento analogo, si mostra che tutti gli insiemi **numerabilmente infiniti** sono misurabili e hanno misura nulla.

[3] Si noti che μ è una *funzione di insieme*, ovvero associa ad un insieme un valore numerico che in questo caso rappresenta la sua misura.

B.2 Funzioni misurabili

Definizione B.4. *Una funzione $f(x)$ avente per dominio il sottinsieme $\mathcal{E} \subseteq \mathbb{R}$ è* ***Lebesgue-misurabile*** *se è verificata una delle seguenti condizioni equivalenti:*

1. $\forall M > 0 \quad \mathcal{E}_M = \{x \in \mathcal{E} : |f(x)| > M\}$ *è* $\mathscr{L}$*-misurabile;*
2. $\forall M > 0 \quad \mathcal{E}_M = \{x \in \mathcal{E} : |f(x)| \geq M\}$ *è* $\mathscr{L}$*-misurabile;*
3. $\forall M > 0 \quad \mathcal{E}_M = \{x \in \mathcal{E} : |f(x)| < M\}$ *è* $\mathscr{L}$*-misurabile;*
4. $\forall M > 0 \quad \mathcal{E}_M = \{x \in \mathcal{E} : |f(x)| \leq M\}$ *è* $\mathscr{L}$*-misurabile.*

Un esempio di funzione non misurabile può essere costruito tramite l'**insieme di Vitali**, che rappresenta un sottoinsieme della retta reale non misurabile da nessuna misura positiva, numerabilmente additiva e invariante per traslazioni. Sapendo questo, una funzione non misurabile è

$$\mathbb{1}_V(x) = \begin{cases} 1 & x \in V \\ 0 & \text{altrove} \end{cases} \tag{B.7}$$

Una definizione alternativa di misurabilità di una funzione è data dalla seguente

Definizione B.5. *f è $\mathscr{L}$-misurabile $\iff$ $f^{-1}(\mathcal{E})$ è $\mathscr{L}$-misurabile $\forall$ $\mathcal{E}$ misurabile.*

Non è difficile convincersi della totale equivalenza fra le due definizioni. Infatti, se vale la Definizione B.5 basta scegliere $\mathcal{E}$ come un intervallo per ritrovare la Definizione B.4. Il viceversa è faccenda più delicata (ma non troppo) che deliberatamente verrà lasciato da parte. Infine,

Definizione B.6. *Due funzioni f e g , $\mathscr{L}$-misurabili su uno stesso insieme, sono* ***$\mathscr{L}$-equivalenti*** *se $\mu(\{x \in \mathcal{E} \;:\; f(x) \neq g(x)\}) = 0$, ovvero se differiscono solo su un insieme di misura nulla. Si scriverà in questo caso $f(x) \overset{L}{\equiv} g(x)$.*

Ad esempio, la funzione di *Dirichlet* è $\mathscr{L}$-equivalente alla funzione identicamente uguale a 0 sull'intervallo $[0, 1]$.

B.3 Integrale di Lebesgue

La differenza sostanziale tra l'integrale di Riemann e l'integrale di Lebesgue consiste nell'approccio al partizionamento dell'intervallo di interesse. Nel caso di Lebesgue, si partiziona il codominio della funzione integranda, prendendo poi le preimmagini degli intervalli consideratisi[4]. A questo punto, definita

$$\sigma_n = \sum_{i=1}^{n} y_i \mu(f^{-1}(\mathcal{E}_i)) \tag{B.8}$$

come successione numerica approssimante il valore dell'integrale, si può definire l'**integrale di Lebesgue** come

[4]Tali preimmagini sono sempre misurabili per definizione, se si lavora con funzioni $\mathscr{L}$-misurabili. Il problema nella teoria di Lebesgue diventa allora capire quali funzioni siano misurabili e, soprattutto, che distanza ci sia tra la classe delle funzioni $\mathscr{L}$-misurabili e la classe delle funzioni continue.

Definizione B.7. $\int_{\mathcal{E}} f(x)\,\mathrm{d}x = \lim_{n,\infty} \sigma_n$

Questa definizione, apparentemente molto simile a quella di Riemann, permette una maggiore elasticità sulle funzioni che è possibile integrare e, in particolare, sulle caratteristiche dei domini di integrazione. Nel caso di Lebesgue, viene meno la necessità che la funzione sia continua quasi ovunque affinchè sia possibile integrarla; viene meno la necessità che il dominio sia un insieme connesso, così che insiemi *discreti* come $\mathbb{N}$ o $\mathbb{Q}$ diventano insiemi su cui è possibile integrare.

B.3.1 Proprietà dell'integrale di Lebesgue

Di seguito sono listate senza dimostrazione alcune **proprietà** dell'integrale:

1. se $\mathcal{E} = [a, b]$ allora $\int_{\mathcal{E},R} f(x)\,\mathrm{d}x = \int_{\mathcal{E},L} f(x)\,\mathrm{d}x$, ovvero se una funzione è Riemann-integrabile su un intervallo chiuso e limitato, allora è Lebesgue-integrabile sullo stesso insieme e gli integrali coincidono;

2. esistono funzioni L-integrabili ma non R-integrabili, come ad esempio la **funzione di Dirichlet**;

3. $\int_{\mathcal{E}_1 \cup \mathcal{E}_2} f(x)\,\mathrm{d}x = \int_{\mathcal{E}_1} f(x)\,\mathrm{d}x + \int_{\mathcal{E}_2} f(x)\,\mathrm{d}x - \int_{\mathcal{E}_1 \cap \mathcal{E}_2} f(x)\,\mathrm{d}x$ (**decomposizione del dominio**);

4. $\int_{\mathcal{E}} f(x)\,\mathrm{d}x = f(\hat{x})\mu(\mathcal{E})$ per qualche $\hat{x} \in \mathrm{dom} f$ (**media integrale**);

5. se f è limitata, sia $M > 0$ tale che $|f| \leq M$, allora $|\int_{\mathcal{E}} f(x)\,\mathrm{d}x| \leq M\mu(\mathcal{E})$.

Si dimostri ora il seguente teorema sulle funzioni $\mathscr{L}$-equivalenti.

Teorema B.1. *Date due funzioni $\mathscr{L}$-misurabili su uno stesso insieme $\mathcal{E}$, se $f \overset{L}{\equiv} g$, allora*

$$\int_{\mathcal{E}} f(x)\,\mathrm{d}x = \int_{\mathcal{E}} g(x)\,\mathrm{d}x\,.$$

Dimostrazione. Sia $\mathcal{F} = \{x \in \mathcal{E} : f(x) \neq g(x)\}$ tale che, per ipotesi, $\mu(\mathcal{F}) = 0$. Si può scrivere $\mathcal{E} = \mathcal{F} \cup (\mathcal{E} - \mathcal{F})$. Allora $\int_{\mathcal{E}} f(x)\,\mathrm{d}x = \int_{\mathcal{F}} f(x)\,\mathrm{d}x + \int_{(\mathcal{E}-\mathcal{F})} f(x)\,\mathrm{d}x = (\star)$. Visto che le due funzioni coincidono su $\mathcal{E} - \mathcal{F}$ e inoltre $\int_{\mathcal{F}} f(x)\,\mathrm{d}x = \int_{\mathcal{F}} g(x)\,\mathrm{d}x = 0$, allora $(\star) = \int_{(\mathcal{E}-\mathcal{F})} g(x)\,\mathrm{d}x + \int_{\mathcal{F}} g(x)\,\mathrm{d}x = \int_{\mathcal{E}} g(x)\,\mathrm{d}x$. □

Questo teorema permette di calcolare con semplicità molti integrali apparentemente complicati. Ad esempio, poiché la funzione di Dirichlet è $\mathscr{L}$-equivalente alla funzione identicamente uguale a 0 in $[0, 1]$, allora essa ha integrale nullo.

B.3.2 Proprietà di convergenza sotto il segno di integrale

Un grande vantaggio nell'utilizzo dell'integrale di Lebesgue sta nel fatto che le condizioni sufficienti affinché valga la formula di passaggio al limite espressa in (B.1) sono più leggere di quelle di cui necessita l'integrale di Riemann. Per quest'ultimo, infatti, vale il seguente teorema:

Teorema B.2. *Data una successione di funzioni $f_n(x)$ tale che $f_n(x) \rightrightarrows f(x)$ su un insieme a misura finita, allora vale* (B.1).

Dimostrazione. $0 \leq \left|\int_{\mathcal{E}}(f_n(x) - f(x))\,\mathrm{d}x\right| \leq \int_{\mathcal{E}}|f_n(x) - f(x)|\,\mathrm{d}x$. Sfruttando le ipotesi di convergenza uniforme e la finitezza della misura di $\mathcal{E}$, risulta $|f_n(x) - f(x)| \leq \frac{\epsilon}{\mu(\mathcal{E})}\ \forall n > n_\epsilon\ \forall x \in \mathcal{E}$. Da questo, segue la tesi sfruttando la proprietà 5. □

Come si osserva, le ipotesi in cui vale il passaggio al limite sotto il segno di integrale sono abbastanza forti: viene richiesta sia la convergenza uniforme della successione di funzioni, sia la finitezza della misura dell'insieme su cui si ha tale tipo di convergenza.
L'**integrale di Lebesgue** permette di alleggerire fortemente tali ipotesi. Si introduce il concetto di **convergenza quasi ovunque**:

Definizione B.8. $f_n \to f$ *q.o.* $\iff$ *la successione converge puntualmente q.o.*

È immediato verificare che se una successione converge puntualmente su un insieme, allora converge quasi ovunque su quell'insieme. Non è in generale vero il viceversa. A questo punto, si può enunciare il seguente

Teorema B.3 (Della convergenza dominata, o di Lebesgue). *Sia $\{f_n\}$ una successione di funzioni $\mathscr{L}$-misurabili su $\mathcal{E}$ $\mathscr{L}$-misurabile, convergente q.o. in $\mathcal{E}$ a f. Se esiste $\varphi(x)$ $\mathscr{L}$-misurabile tale che $|f_n(x)| \leq |\varphi(x)|$ q.o. definitivamente*[a] *in $\mathcal{E}$ e $\int_{\mathcal{E}} \varphi(x)\,\mathrm{d}x < \infty$, si ha che*

$$\lim_{n,\infty} \int_{\mathcal{E}} f_n(x)\,\mathrm{d}x = \int_{\mathcal{E}} \lim_{n,\infty} f_n(x)\,\mathrm{d}x\,.$$

[a] Ovvero $\forall n > n_0 \in \mathbb{N}$, per ogni n maggiore di un certo n_0 finito.

Così, ad esempio, la successione di funzioni $f_n(x) = \frac{e^{-nx}}{1+x^2}$ è dominata da $\varphi(x) = \frac{1}{1+x^2}$. È, dunque, immediato stabilire che $\lim \int_{\mathbb{R}} f_n(x)\,\mathrm{d}x = 0$.

Appendice C

Simboli e abbreviazioni

$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$	insiemi numerici standard
SVL	spazio vettoriale lineare
$\mathbb{K}^n$	spazio dei vettori a n componenti in $\mathbb{K}$
$\mathbb{C}$	insieme dei numeri complessi
$\mathbb{K}$	un generico campo
$\overline{z}$, z^*	complesso coniugato di $z \in \mathbb{C}$
$A^\dagger$	aggiunto dell'operatore A
$\hat{f}$	Trasformata di Fourier di f
δ_{ij}	delta di Kronecker
$\mu(\mathcal{E})$	misura di Lebesgue dell'insieme $\mathcal{E}$
$\overset{L}{\equiv}$	Lebesgue-equivalente, o uguaglianza q.o.
$\int_{\mathbb{R}}$	integrale tra $-\infty$ e $+\infty$
$\mathscr{L}^p$	spazio delle funzioni f Lebesgue misurabili a potenza p-esima integrabile
ℓ^p	spazio delle successioni a somma p-assolutamente convergente
H	spazio di Hilbert
$\langle a, b \rangle$	prodotto scalare tra a e b
$\lVert\cdot\rVert$	norma

$\lVert\cdot\rVert_p$	norma p-esima nel senso di $\mathscr{L}^p$, $1 \leq p \leq +\infty$
$\rightrightarrows$	converge uniformemente a
$\xrightarrow[n\to\infty]{\tau_\Gamma}$	converge τ_Γ per $n \to \infty$ a
$\xrightarrow[\alpha\to 0]{\mathbb{C}}$	converge in $\mathbb{C}$ per $\alpha \to 0$ a
$\chi\text{-}\lim\limits_{x,x_0} \psi(x)$	il χ-limite di $\psi(x)$ per $x \to x_0$
q.o.	quasi ovunque
: , | , t.c.	tale (tali) che
$:=$	definito come, uguale per definizione a
$\Longrightarrow$, $\Rightarrow$	implica, si ha che
$\Longleftrightarrow$	se e solo se
$\therefore$	ovvero
$\lightning$	contraddizione (dimostrazioni per assurdo)
$\mathscr{S}(\mathbb{R})$	spazio di Schwarz, ovvero delle funzioni reali a decrescrenza rapida
$\mathscr{S}'(\mathbb{R})$	spazio delle distribuzioni temperate
$\mathcal{D}(\mathbb{R})$	spazio di Dirichlet, ovvero delle funzioni $C^\infty(\mathbb{R})$ a supporto compatto
$\mathcal{D}'(\mathbb{R})$	spazio delle distribuzioni
$\text{supp}\{f\}$	supporto di f, ovvero insieme dei punti in cui $f \neq 0$
$\delta(x)$	delta di Dirac
$f\vert_{\mathbb{R}}$, $f\vert_{-\infty}^{+\infty}$	l'espressione f è valutata come $\lim\limits_{x,+\infty} f(x) - \lim\limits_{x,-\infty} f(x)$
$\Im\{z\}$	parte immaginaria di $z \in \mathbb{C}$
$\Re\{z\}$	parte reale di $z \in \mathbb{C}$
$\mathbb{K}^{i\times j}$	matrice $i \times j$ a entrate in $\mathbb{K}$

Indice analitico